

**Univerzita Karlova v Praze**

**Filozofická fakulta**

Fonetický ústav

# **Bakalářská práce**

Alžběta Růžicková

**Strategie maskování hlasu u českých mluvčích**

Voice Disguise Strategies of Czech Speakers

Praha 2016

Vedoucí práce: Doc. Mgr. Radek Skarnitzl, Ph.D.

Děkuji vedoucímu práce Doc. Mgr. Radku Skarnitzlovi, Ph.D. za odborné vedení, cenné rady a vstřícnost. Mé poděkování patří také respondentům, kteří se zúčastnili nahrávání a percepčních testů.

Prohlašuji, že jsem bakalářskou práci vypracovala samostatně, že jsem řádně citovala všechny použité prameny a literaturu a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.

V Praze, dne 24. srpna 2016

.....

Jméno a příjmení

## **Abstrakt**

Ve forenzní praxi, pokud je zapotřebí identifikovat pachatele na základě nahrávky jeho řečového projevu, se uplatňuje fonetická identifikace mluvčího. Pokud si je pachatel vědom, že jeho projev může být nahráván, např. v případech, jakými jsou únosy či vydírání, může se pokusit o maskování svého hlasu, aby tak zkomplikoval proces identifikace. Tato práce mapuje strategie, jež uplatňují čeští mluvčí, pokud mají libovolným způsobem maskovat svůj hlas. Byla provedena poslechová analýza maskovaných projevů 80 mluvčích a u 15 z nich, kteří využili komplexnějších způsobů maskování hlasu, byla následně provedena také akustická analýza základní frekvence, vokálních formantů a mluvního a artikulačního tempa. Pro tyto mluvčí byly také sestaveny percepční testy ověřující míru odhalitelnosti modifikace řečového projevu a dopad zvolených způsobů maskování hlasu na identifikovatelnost mluvčích při rekognici.

## **Klíčová slova**

Fonetická identifikace mluvčího, maskování hlasu, forenzní fonetika, rekognice podle hlasu

## **Abstract**

In forensic practice, when there is the need to identify a speaker using a recording of his speech, forensic voice identification is applied. If the perpetrator is aware of the option of their speech being recorded, for instance in cases such as abduction or blackmail, they can attempt voice disguise in order to thwart voice identification process. This thesis surveys strategies used by Czech speakers, when they attempt to disguise their voices using any means of their choice. Auditory analysis of 80 speakers' disguised voices was performed. 15 of them, who used more complex strategies for disguising their voices, acoustic analysis of fundamental frequency, formant values and speech and articulatory rate was performed. For those speakers, perceptual tests were carried out, to examine the degree of difficulty in uncovering modifications of one's speech and the effect of the used disguise strategies on speaker identifiability in a voice line-up.

## **Key words**

Forensic voice identification, voice disguise, forensic phonetics, voice line-up

# OBSAH

1.	Identifikace mluvčího ve forenzně fonetické praxi .....	8
1.1.	Profilování mluvčího .....	8
1.2.	Sporná a srovnávací nahrávka .....	9
1.3.	Rekognice svědkem .....	10
2.	Idiosynkratické prvky řeči .....	12
2.1.	Vokálníkové formanty .....	12
2.2.	Základní frekvence .....	13
2.3.	Barva hlasu, spektrální sklon .....	14
2.4.	Temporální ukazatele .....	16
2.5.	Dysfluence .....	17
2.6.	Přízvuk, dialekt a cizí jazyk .....	17
3.	Modifikace hlasu v závislosti na okolnostech .....	19
4.	Maskování hlasu .....	21
4.1.	Změny fonace .....	22
4.2.	Napodobování dialektu a cizineckého přízvuku .....	25
4.3.	Cizí objekt v ústech .....	26
4.4.	Elektronické maskování hlasu .....	27
5.	Strategie maskování hlasu u českých mluvčích .....	28
6.	Metoda .....	29
6.1.	Materiál .....	29
6.2.	Procedura .....	29
6.2.1.	Poslechová analýza .....	29
6.2.2.	Akustická analýza .....	30
6.2.3.	Percepční testy .....	30
7.	Výsledky .....	33
7.1.	Poslechová analýza .....	33
7.2.	Akustická analýza .....	37
7.2.1.	Základní frekvence .....	37
7.2.2.	Vokálníkové formanty .....	39
7.2.3.	Tempo .....	40
7.3.	Percepční testy .....	41
7.3.1.	Percepční test č. 1 .....	41
7.3.2.	Percepční test č. 2 .....	44
8.	Souhrn výsledků pro jednotlivé mluvčí .....	47
8.1.	Mluvčí DRID .....	47
8.2.	Mluvčí JELS .....	47
8.3.	Mluvčí KAND .....	48
8.4.	Mluvčí KATR .....	49
8.5.	Mluvčí KREM .....	49

8.6.	Mluvčí KUCK.....	50
8.7.	Mluvčí MAKN .....	51
8.8.	Mluvčí MAND .....	51
8.9.	Mluvčí OLET.....	52
8.10.	Mluvčí ROLK.....	53
8.11.	Mluvčí SJAN.....	53
8.12.	Mluvčí VEDR .....	54
8.13.	Mluvčí VLAJ .....	54
8.14.	Mluvčí ZALT .....	55
8.15.	Mluvčí ZENT.....	56
9.	Diskuse .....	57
	Reference.....	61
	Přílohy .....	63
	Příloha 1 – předčítané texty .....	63
	Příloha 2 – výsledky akustické analýzy vokálních formantů při maskovaných a nemaskovaných projevech mluvčích a vyčíslení p-hodnoty pomocí dvouvýběrového t-testu .....	64

## **1. Identifikace mluvčího ve forenzně fonetické praxi**

Každý mluvčí se vyznačuje specifickým řečovým projevem. Ačkoli se lidé, resp. příslušníci jedné jazykové komunity, vzájemně dorozumívají pomocí jednoho ustáleného kódu, tj. jazyka, projev každého z nich vykazuje prvky, jež jsou pro daného mluvčího charakteristické – tzv. idiosynkratické jevy.

Na základě těchto jevů lze provádět rozpoznávání mluvčích (speaker recognition). Tento koncept zahrnuje verifikaci (speaker verification) a identifikaci mluvčích (speaker identification). Zatímco při verifikačním procesu mluvčí chce být rozpoznán (např. v případě umožnění přístupu na základě ověření hlasu), tj. jedná se o ověření totožnosti známého jedince, při identifikaci jde o rozpoznání neznámého mluvčího pomocí analýzy hlasu. Fonetická identifikace mluvčího se uplatňuje např. ve forenzně fonetické praxi, je-li zapotřebí určit pachatele na základě jeho řečového projevu.

### **1.1. Profilování mluvčího**

Pokud je k dispozici sporná nahrávka (evidence tape) projevu pachatele pořízená při činu (např. telefonní odposlech), lze na základě její analýzy provést profilování mluvčího a určit sociolingvistické klasifikační parametry (pohlaví, věk, vzdělání, geografickou oblast původu, mateřský jazyk) a také zaznamenat nápadné, ze statistického hlediska neobvyklé idiosynkratické rysy v řečovém projevu, jakými mohou být např. řečové vady, užívání neobvyklých lexikálních jednotek aj. (Skarnitzl, 2014).

Určení pohlaví bývá obvykle jednoznačné díky odlišným fyziologickým vlastnostem mužů a žen; dospělé ženy mají v průměru o 15 % kratší vokální trakt než muži (Fitch & Giedd, 1999; citováno v Skarnitzl, 2014), což se odráží ve vyšší frekvenci formantů, a blanitá část glottis je u žen asi o 60 % kratší (Titze, 1989; citováno v Skarnitzl, 2014), což vede k vyšší základní frekvenci kmitání hlasivek. Fyziologické rozdíly ale mohou být do jisté míry potlačeny sociokulturními faktory (Skarnitzl, 2014). Věk bývá ve forenzní praxi odhadován přibližně v desetiletém rámci, jeho určování však mnohdy nebývá příliš spolehlivé. Probíhají výzkumy s cílem nalezení korelace mezi akustickými parametry řečového projevu a tělesnou výškou, ta však zřejmě není příliš vysoká, stejně jako je tomu v případě indexu tělesné hmotnosti BMI (Skarnitzl, 2014).

Profilování a identifikace mluvčího na základě hlasu nebývají zcela jednoznačné. Hollien (2002) zmiňuje myšlenku „otisku hlasu“ (voiceprint) jakožto analogie otisku



prstu, kdy je hlasový projev nazírán jakožto něco charakteristického, čím se snadno a jednoznačně odliší jeden jedinec od jiných, upozorňuje však, že tento přístup je zcestný, neboť lidský hlas je velmi variabilní a podoba řečového projevu jedince je závislá na mnoha okolnostech. Vokální trakt každého mluvčího je sice unikátní, ale jeho stavba neurčuje jednoznačně hodnoty konkrétních řečových parametrů; žádné parametry, které by s jistotou odlišily jednoho mluvčího od všech ostatních, nebyly nalezeny (Skarnitzl, 2014).

## **1.2. Sporná a srovnávací nahrávka**

Jestliže byl zadržen podezřelý, jenž by mohl být mluvčím zaznamenaným na sporné nahrávce, lze pořídit tzv. srovnávací nahrávku (exemplar tape) hlasu podezřelého a provést porovnání se spornou nahrávkou. Hollien (2002) doporučuje získat tři různé typy vzorků hlasu: nepřipravený projev (pod co nejmenším stresem, vyšetřovatel by tedy měl zavést hovor na neutrální témata, cílem je získat co nejpřirozenější projev), opakování (části) textu zachyceného na sporné nahrávce a čtení úseků z novin či časopisu. Cílem je získání co možná nejširšího repertoáru řeči subjektu, tímto způsobem je také někdy možné překazit pokusy o maskování hlasu. Opakování textu obsaženého na sporné nahrávce také umožňuje porovnat totožné řetězce segmentů.

Na rozdíl od srovnávací nahrávky bohužel není možné ovlivnit podobu nahrávky sporné, jejíž kvalita může být degradována okolním ruchem či vlastnostmi nahrávacího zařízení (channel distortions) nebo změnami charakteru řečového projevu mluvčího v závislosti na jeho momentálním stavu, např. stresu či nemoci (speech distortions). Existuje také možnost, že pachatel maskuje svůj hlas. Tyto okolnosti komplikují proces identifikace mluvčího (viz dále).

K identifikaci mluvčího lze přistupovat několika způsoby. Laik i profesionál ve fonetické oblasti může uplatnit poslechový přístup, kdy pečlivě poslechově analyzuje řečové vzorky a formuluje závěry na základě svého percepčního dojmu. Školení fonetici mohou zaujmout přístup akustický a provádět identifikaci mluvčího na základě akustické analýzy řečového signálu. Třetí možností je přístup automatický, při kterém jsou k identifikaci mluvčího využívány specializované počítačové aplikace.

Hollien (2002) představuje poslechovou metodu srovnávání sporné a srovnávací nahrávky, kdy je podobnost zachycených hlasů vyhodnocována na škále až pro dvacet faktorů. Výsledky pak vedou k rozhodnutí, zda se jedná o téhož mluvčího, s určitým konfidenčním intervalem. Vyhodnocení jsou stanovována pro následující kategorie

parametrů: (1) výška, (2) barva hlasu, (3) variování intenzity, (4) dialekt, (5) segmentální faktory, (6) prozodie, (7) případné řečové vady. Výzkumník poslouchá vzorky v párech (sporná + srovnávací nahrávka) a porovnává je (kontrastuje neznámého mluvčího – pachatele se známým – podezřelým). Toto porovnávání probíhá pro každý faktor samostatně, vzorky jsou opakovány, dokud výzkumník nerozhodne o míře podobnosti a nevyznačí tuto míru na škále pro daný jev. Tyto škály jsou desetibodové, přičemž hodnocení 0-3 znamenají, že se zřejmě jedná o dva různé mluvčí, naopak hodnoty 7-10 značí, že jde pravděpodobně o téhož mluvčího. Hodnoty 4-6 znamenají možnost shody, hlasy se nápadně neliší a mohly by být produkovány týměž jedincem.

### **1.3. Rekognice svědkem**

Existuje-li svědek, který zaslechl hlas pachatele, jež však neviděl, naskýtá se možnost provést rekognici podle hlasu. Svědkovi není přímo představován hlas podezřelého, kdy by svědkovým úkolem bylo určit, zda se jedná o pachatele, či nikoliv. Nahrávka projevu podezřelého je svědkovi prezentována spolu s hlasy několika figurantů (stejně jako je tomu při rekognici pachatele očitým svědkem, kdy je před svědka předvedeno několik figurantů společně s podezřelým, načež svědek z těchto osob vybírá pachatele). Svědek je požádán, aby identifikoval, který z prezentovaných hlasů patří pachateli. Hollien (2002) však uvádí, že tento způsob identifikace není příliš robustní, neboť svědectví jednotlivce může být nespolehlivé, připouští však, že může do jisté míry napomoci vyšetřování.

Při rekognici podle hlasu může dojít k nesprávnému určení kvůli očekávání svědka – svědci mají tendenci předpokládat, že hlas pachatele se při rekognici objeví; pokud se tak nestane, často je označen některý figurant. Ukázalo se, že velmi málo posluchačů zvolí možnost, že hlas podezřelého nezazněl, i v případě, že byli zpraveni o tom, že taková možnost existuje (Hollien, 2002).

Hollien doporučuje při rekognici využít pěti až osmi figurantů s hlasy různého charakteru: jeden či více mohou být podobné hlasu podezřelého, jeden či více by se od něj měly výrazněji lišit a zbytek by měl být „něco mezi“. Vzorky prezentované svědkovi by měly být uniformní – nahrané ve stejném prostředí, s podobným textovým obsahem atd.

Schopnost svědka identifikovat hlas pachatele je ovlivňována různými faktory. Jedním z nich je známost svědka a pachatele – výzkumy prokázaly, že známé hlasy jsou rozeznávány lépe než neznámé (Hollien, 2002). Na přesnost identifikace má vliv také

doba, která uběhla mezi zaslechnutím pachatele svědkem a rekognicí. Výzkum McGehee (1944; citováno v Hollien, 2002) ukázal postupný časový sestup schopnosti rozpoznat hlas neznámého mluvčího; po jednom dni byla úspěšnost identifikace asi 83 %, tato úroveň byla více méně udržena asi po týden. První propad se objevil po dvou týdnech (68% úspěšnost), další po třech měsících (35 %). Po pěti měsících byli respondenti schopni správně identifikovat mluvčího, kterého prve slyšeli, asi ve 13 % případů (tj. pod úrovní náhody). „Dosud není možné přesně specifikovat, jak bude vypadat křivka [časového sestupu hlasové paměti] za všech okolností nebo co může být očekáváno u konkrétního jedince. Výsledek může ovlivnit příliš mnoho proměnných“ (Hollien, 2002: 31).

## 2. Idiosynkratické prvky řeči

V této kapitole jsou předloženy jevy rozeznatelné v řečovém projevu jedince, které mohou napomoci jeho identifikaci. Jak již bylo zmíněno, žádný z těchto faktorů sám o sobě neposkytuje dostatečnou oporu pro jednoznačné určení identity mluvčího, pomocí jejich kombinace však lze dospět k poměrně spolehlivým závěrům.

Na některé z těchto jevů se mohou při poslechu zaměřovat i laici v oblasti fonetiky, pokud jsou požádáni o podání svědectví; např. na přízvuk, dialekt, výšku a variabilitu základní frekvence či tempo řeči. Podrobněji jsou předložené faktory analyzovány odborníky ve forenzní fonetické praxi, kteří provádí jejich akustickou analýzu a závěry formulují na základě jejich výsledků.

### 2.1. Vokální formanty

Faktorem často využívaným při fonetické identifikaci mluvčího jsou vokální formanty, tj. „rezonanční frekvence vzduchového sloupce ve vokálním traktu“ (Skarnitzl, 2014: 21). V řečovém signálu se tyto frekvence, resp. frekvenční pásma projevují vyšší intenzitou. Hodnoty vokálních formantů se mění podle momentálního nastavení artikulačních orgánů. Kvalitu vokálů určují zejména F1 a F2 (tj. první a druhý formant), částečně je na kvalitě vokálů závislý také F3. Vyšší formanty, F4 a F5, závisí spíše na konkrétním mluvčím, resp. na vlastnostech jeho vokálního traktu, nemění se v závislosti na kvalitě vokálu. Vokální formanty obecně jsou tradičně považovány za jedny z nejužitečnějších parametrů pro identifikaci mluvčího (Hollien, 2002), neboť do jisté míry odrážejí jak fyziologické vlastnosti mluvčího, tak i různé faktory sociolingvistické povahy (regionální původ, věk aj.). Skarnitzl (2014) však podotýká, že ačkoliv by se mohlo zdát, že vyšší formanty budou díky své neměnnosti během artikulace pro forenzní fonetiku přínosnější, jejich nevýhoda spočívá v tom, že v reálných forenzních podmínkách se může stát, že v signálu nebudou dobře viditelné nebo ani přítomné.

Při fonetické identifikaci mluvčího lze vokální formanty využít několika způsoby. Nejjednodušší metodou je analýza průměrných frekvencí formantů v jednotlivých vokálech. Měření lze provést v artikulačním cíli vokálu (tj. v místě, kde formanty dosáhly stabilního průběhu), případně v prostřední části vokálu (zpravidla v prostřední třetině či v úseku mezi 25 a 75 % trvání vokálu).

Podle Skarnitzla (2014) mohou první dva formanty přispět k odlišení mluvčích; zcela odlišné hodnoty pro různé mluvčí nejsou pravidlem, ve forenzních nahrávkách ale

mohou frekvence měřené v prostřední části vokálů poskytnout užitečné informace směřující k potvrzení či vyloučení totožnosti mluvčích.

Dalším způsobem využití vokalických formantů při identifikaci mluvčího je pozorování dynamického průběhu formantových hodnot, tedy porovnávání formantových trajektorií. Namísto hodnoty v jednom bodě či úseku vokálu jsou v tomto případě vyčísleny hodnoty formantů v několika ekvidistantních bodech. Greisbach, Esser & Weinstock (1995; citováno v Skarnitzl, 2014) experimentálně prokázali, že použití pěti hodnot extrahovaných z monoftongů a diftongů může vést až k třicetiprocentnímu nárůstu úspěšnosti při identifikaci mluvčích ve srovnání s využitím samotné střední hodnoty.

Skarnitzl (2014) srovnával využitelnost formantových trajektorií pro identifikaci mluvčího v češtině a angličtině a dospěl k závěru, že pro odlišnost vokalických systémů těchto jazyků je v případě této metody prostor pro idiosynkratičnost v češtině ve srovnání s angličtinou omezený. Výhodu naopak spatřuje ve skutečnosti, že v češtině není třeba se omezovat pouze na přízvučné vokály jako v angličtině, nepřízvučné vokály naopak zřejmě mohou poskytovat o mluvčím více informací.

Třetí metoda využívá dlouhodobých vlastností formantů, jedná se o dlouhodobé formantové distribuce (long-term formant distribution, LTF). Jejich hodnoty jsou vyčíslovány pro konkrétní formanty daného mluvčího: jsou sloučeny všechny frekvence detekované pro daný formant do intervalů o šířce 25 Hz a následně je vytvořen jejich histogram. Jde o průměrnou distribuci sledovaného formantu bez ohledu na identitu vokálů. Formantové distribuce jsou pro jednoho mluvčího stabilní a různé mluvčí od sebe významně odlišují, jako nejužitečnější pro identifikaci mluvčího i nejrobustnější vůči zkreslení se jeví LTF3, tj. dlouhodobá distribuce třetího formantu (Skarnitzl, 2014).

## **2.2. Základní frekvence**

Kromě frekvencí vokalických formantů lze z řečového signálu extrahovat také základní frekvenci  $f_0$ , která je akustickým korelátem frekvence kmitání hlasivek při produkci znělých hlásek. Pro účely forenzně fonetické analýzy je tento akustický ukazatel užitečnější než jeho percepční korelát, kterým je vnímaná výška hlasu; „dojem vysokého či nízkého hlasu může být tvořen více jeho vlastnostmi, kromě základní frekvence především vokalickými formanty, a tyto vlastnosti od sebe můžeme pomocí akustické analýzy oddělit“ (Skarnitzl, 2014: 49).

Jedním z ukazatelů, který by měl pokud možno charakterizovat mluvčího jako jedince a zároveň jej vymezit oproti ostatním, je hodnota střední hlasové frekvence, tedy střední hodnota  $f_0$ . Tradičně se za ni považuje aritmetický průměr, případně medián či modus hodnot základní frekvence řečového signálu mluvčího. V některých jazycích jsou pro tento parametr dostupné populační statistiky (pro češtinu však doposud bohužel nikoliv), při konkrétním forenzním případě je pak možné porovnat, nakolik je střední hlasová frekvence sledovaného mluvčího typická pro danou populaci nebo naopak nakolik vybočuje z průměru.

Střední hlasovou frekvenci může ovlivnit řada faktorů. Braun (1996; citováno v Skarnitzl, 2014) je dělí do tří skupin: na vlivy technické (maskování elektronické i přirozené, brum z elektrické sítě, brum nahrávacího zařízení, množství dostupného řečového materiálu), fyziologické (věk, požití alkoholu či drog, kouření – u kuřáků je v průměru pozorována nižší  $f_0$ ) a psychologické (zejména různé afektivní stavy – zvýšení i vyšší variabilita  $f_0$  při radosti či vzteku, naopak snížení obou ukazatelů např. při smutku; při stresu dojde většinou ke zvýšení  $f_0$ , tato proměnná se však u jednotlivých mluvčích může lišit). Skarnitzl (2014) doplňuje aspekty situační, kam řadí např. denní dobu či mluvní styl, dodává také, že střední hlasová frekvence se liší i v různých jazycích či dokonce dialektech jednoho jazyka, lze tedy hovořit o faktorech kulturních.

Kromě střední hlasové frekvence o identitě mluvčího vypovídá také variabilita jeho základní frekvence, tedy její rozptyl kolem průměru.

Výhodou základní frekvence je její běžná dostupnost i v krátkých řečových vzorcích, kromě toho je robustní vůči zkreslení včetně telefonního přenosu a zároveň vykazuje poměrně vysokou variabilitu mezi mluvčími; značnou nevýhodu však představuje vysoká variabilita intraindividuální, tj. v rámci řečového projevu jednoho mluvčího (Skarnitzl, 2014).

### **2.3. Barva hlasu, spektrální sklon**

Jednou z nejvýraznějších individuálních řečových charakteristik mluvčího je barva hlasu. Jejím akustickým korelátem je spektrální složení, tj. zastoupení, resp. prominence jednotlivých frekvencí v řečovém signálu. Spektrální složení hlasu mluvčího je dáno vlastnostmi a momentálním nastavením vokálního traktu.

Důležitým ukazatelem je spektrální sklon, tj. postupná deklinace intenzity s narůstající frekvencí v řečovém spektru. „Spektrální sklon se obvykle pohybuje kolem 12 dB na

dekádu (tj. pokles amplitudy o 12 dB na desetinásobné frekvenci), záleží však na tónu, hlasitosti a fonačním nastavení konkrétního mluvčího“ (Skarnitzl, 2014: 79).

Obecně platí, že hlasy s relativně velkým zastoupením vyšších frekvencí znějí „ostře“, zatímco čím strmější je spektrální sklon, tím hlas vyznívá „měkčeji“. Spektrální sklon je zpravidla zmenšován při vyšším mluvním úsilí, stejný efekt má také falzet nebo třepená fonace.

„Protože spektrální sklon do velké míry závisí na vlastnostech vokálního traktu mluvčího a na charakteristickém způsobu fonace, je využíván také pro rozpoznávání mluvčího“ (Skarnitzl, 2014: 84).

Ve forenzně fonetické praxi může být měření spektrálního sklonu řečového projevu mluvčího komplikováno nedostatečnou kvalitou či hlasitostí nahrávky (zesílení má za následek aditivní šum v signálu, který zkreslí spektrum). Zásadní dopad na spektrum analyzované řeči může mít také např. prostředí, ve kterém byla nahrávka pořízena, může být přítomen okolní ruch či odraz zvuku od stěn či podlahy, podobu spektra ovlivňují také vlastnosti mikrofону a poloha mluvčího vůči němu. Pokud je k dispozici pouze různými způsoby degradovaný signál (tj. nikoliv nahrávka studiové kvality), je třeba brát v potaz, jaké prvky mohou ovlivňovat tvar spektra a tím i spektrální sklon a z nahrávek pak cíleně vybírat takové části, které jsou co nejméně zkreslené a obsahují minimum ruchů na pozadí (Skarnitzl, 2014). French (1994; citováno v Skarnitzl, 2014) doporučuje analyzovat pouze nahrávky srovnatelné kvality, a pokud takové vzorky nejsou k dispozici, pak jeden z nich záměrně kvalitativně degradovat tak, aby si oba vzorky byly co nejpodobnější.

Na základě změn spektra v průběhu projevu mluvčího lze získat tzv. dlouhodobé průměrné spektrum (long-term average spectrum, LTAS), které poskytuje informace o spektrálním rozložení energie v řečovém signálu během delšího časového úseku. Krátkodobé změny dané segmentálním složením projevu mluvčího jsou potlačeny, výsledné spektrum tedy reflektuje celkové přispění zdroje (hlasivek) i vokálního traktu ke kvalitě hlasu mluvčího a není ovlivněno segmentálními rozdíly v řečovém materiálu. Podle dosavadních výzkumů může LTAS fungovat jako užitečný ukazatel rozdílů mezi mužskými a ženskými hlasy, rozdílů ve věku nebo mezi trénovanými a netrénovanými hlasy (Skarnitzl, 2014). Přestože výhoda dlouhodobých ukazatelů, jakým je i LTAS, spočívá v eliminaci vlivu jednotlivých hlásek, a tyto ukazatele tedy poskytují informace o průměrných hodnotách v řeči daného mluvčího, jsou i ony ovlivňovány různými faktory a mohou značně variovat i v rámci řečového projevu jednoho mluvčího

(Löfqvist 1986; citováno v Skarnitzl, 2014). Při využití LTAS pro forenzní účely je tedy třeba brát v úvahu všechny faktory, které na něj mohou mít vliv.

Kromě dlouhodobého průměrného spektra lze k identifikaci mluvčího využít také vykreslení krátkodobého průměrného spektra, tj. pouze krátké části zvukového signálu, typicky jedné hlásky, případně její části. Zde hraje velkou roli identita zkoumaného segmentu, kromě informací o onom segmentu však může tento ukazatel vypovídat také o totožnosti daného mluvčího.

#### **2.4. Temporální ukazatele**

Spektrální složení (včetně základní frekvence a vokálních formantů) určuje výslednou kvalitu hlasu. Tento ukazatel, jak již bylo zmíněno, může být zkreslen v důsledku kvalitativní degradace signálu. Vůči tomuto zkreslení jsou naopak odolné temporální charakteristiky řečového projevu, tj. jevy související s trváním. Zdá se, že jsou také odolné vůči úmyslnému maskování, neboť si jich mluvčí často nejsou vědomi (Eriksson & Wretling, 1997; Wretling & Eriksson, 1998; Dellwo, Ramyeard & Dankovičová, 2009; citováno v Skarnitzl, 2014).

V rámci celého průběhu promluvy lze sledovat globální temporální ukazatele, kterými jsou mluvní a artikulační tempo. V obou případech jsou vyčíslovány počty lingvistických jednotek (hlásek, slabik či slov) za určitou časovou jednotku, avšak zatímco při výpočtu mluvního tempa jsou zahrnuty i pauzy, hezitace a dysfluence objevující se v řečovém projevu, v případě tempa artikulačního jsou započítány pouze plynulé části promluv. Na rozdíl od artikulačního tempa má tedy na tempo mluvní vliv např. frekvence a délka pauz, v nichž se jednotliví mluvčí mohou lišit, rozdíly se však mohou objevovat i u jednoho mluvčího v závislosti na mluvním stylu či okolních podmínkách. Pro forenzní účely je tedy výhodnějším ukazatelem artikulační tempo, které je v rámci projevu jednoho mluvčího stabilnější (Skarnitzl, 2014).

V průběhu promluvy dochází ke změnám tempa řeči, např. závěrovému zpomalování či akomodaci vůči komunikačnímu partnerovi při dialogu. Kromě globálních temporálních ukazatelů jsou tedy také sledovány temporální ukazatele lokální, např. lokální artikulační tempo, které je využíváno pro sledování změn temporálních charakteristik v okolí prozodických předělů (Volín, 2009b; citováno v Skarnitzl, 2014).

Případná využitelnost pro forenzní praxi je testována také pro rytmické ukazatele, které jsou obvykle založeny na podílu, případně variabilitě trvání vokálních a konsonantických intervalů v řeči; primárně je tento parametr využíván pro popis jazyků



a dialektů, studie však ukázaly značnou variabilitu mezi mluvčími. Nejčastěji využívanými rytmickými ukazateli jsou např. %V (podíl vokálních intervalů v řeči) či  $\Delta V$  a  $\Delta C$  (směrodatná odchylka trvání vokálních a konsonantických intervalů v řeči).

## **2.5. Dysfluence**

V řečovém projevu mluvčího, zejména jedná-li se o projev nepřipravený, se mohou v různé míře objevovat dysfluence, tj. neplynulosti v řečovém kontinuu: vyplněné nebo tiché pauzy, opakování celých slov či jejich částí nebo naopak celých frází, „falešné začátky“ (pokud se mluvčí rozhodne přerušit započatou promluvu a pronést jinou) a prodloužení konsonantů či vokálů.

Podle výzkumu Braun & Resin (2015) se mluvčí různí v produkci jednotlivých typů dysfluencí, přičemž u daného mluvčího většinou převažují jeden nebo dva typy, avšak žádný z typů výrazně nedominuje u všech mluvčích. Individuální vzorce dysfluentního chování jsou v rámci projevu jednoho mluvčího variabilní, ne však do té míry, aby byly snadno zaměnitelné se vzorci jiných mluvčích. Autorky tedy považují počet a distribuci těchto prvků za slibný jev pro identifikaci mluvčích. Zároveň však podotýkají, že pro spolehlivé posouzení výskytu dysfluencí je zapotřebí více materiálu, než je v mnoha forenzních případech k dispozici, proto tento přístup nelze využívat samostatně, může ale podpořit identifikaci mluvčího v kombinaci s dalšími metodami.

## **2.6. Přízvuk, dialekt a cizí jazyk**

Jednotliví mluvčí se mohou navzájem lišit regionálním či sociálním zabarvením svého řečového projevu; ten může vykazovat specifický dialekt (regionální varietu jazyka), sociolekt (společenskou varietu jazyka) či idiolekt (varietu jazyka specifickou pro daného mluvčího, např. co se týče volby lexikálních prostředků). Pokud je projev mluvčího ovlivněn jiným jazykem (např. jeho mateřštinou), může v něm být patrný cizinecký přízvuk.

Dialekt zaznamenaný v řečových vzorcích, tj. na sporné i srovnávací nahrávce, může být využit při vyhodnocování identity mluvčího. Je analyzován suprasegmentálně i segmentálně a také co se týče jeho výraznosti. Je určováno, zda je dialekt dán regionálně (v rámci daného jazyka) či zda je výsledkem dopadu cizího jazyka, např. mateřského jazyka mluvčího. Zjištění, že oba řečové vzorky se vyznačují stejným dialektem, není příliš robustním ukazatelem, že se jedná o téhož mluvčího, pokud se nejedná o dialekt velmi ojedinělý, neboť na nahrávkách mohou být zachyceni různí mluvčí hovořící tímž dialektem, není tedy možné potvrdit totožnost mluvčího pouze na

základě tohoto jevu. Naproti tomu pokud je na každé z nahrávek nalezen odlišný dialekt, je vysoce pravděpodobné, že se jedná o dva různé mluvčí (Hollien, 2002).

V případě, že mluvčí hovoří cizím jazykem, je pro posluchače identifikace mluvčího obtížnější (Schiller & Köster, 1996; citováno v Hollien, 2002).

Hollien (2002) upozorňuje, že ačkoliv dialekt či přízvuk mohou napomoci identifikaci mluvčího, ve většině případů přesnost identifikace snižují, neboť odvádí pozornost posluchače, často zastírají některé idiosynkratické rysy, potlačují či eliminují drobné, avšak užitečné nuance v jazyce a často redukuje počet dostupných segmentálních kontrastů. Na druhou stranu však dosud nebylo experimentálně prokázáno, že by cizí jazyk, dialekt či přízvuk měl negativní vliv na identifikaci mluvčího, pokud je prováděna pomocí vhodných počítačích metod. Zdá se tedy, že závažnější negativní dopad cizího jazyka či dialektu se objevuje při identifikaci mluvčího neprofesionálními posluchači.

### 3. Modifikace hlasu v závislosti na okolnostech

Jednotlivé parametry řečového projevu se mění podle situace a momentálního rozpoložení mluvčího. Jak již bylo řečeno, ukazatele, jakými jsou např. základní frekvence, temporální ukazatele, výskyt dysfluencí v řeči či síla přízvuku či dialektu, jsou ovlivňovány mimo jiné mluvním stylem, tj. liší se např. při připraveném formálním projevu a projevu neformálním a nepřipraveném.

Dalším faktorem ovlivňujícím podobu řečového projevu jsou afektivní stavy – pokud má mluvčí radost či vztek, dojde ke zvýšení  $f_0$  i její variability, oba ukazatele jsou naopak sníženy např. při smutku. Ve forenzních případech se na pořízených nahrávkách často objevuje stres, jehož projevy v řeči se u mluvčích mohou značně lišit (Skarnitzl, 2014). Pro forenzní účely jsou v řečovém projevu také zkoumány indikátory lži.

Byl sledován vztah mezi lží a psychologickým stresem a jejich projevy v lidském hlase. Komerční systémy zaměřující se na detekci akustických korelátů stresu a lži v řeči jsou na trhu přes třicet let, jejich efektivita však nebyla úspěšně ověřena (Harnsberger et al., 2009). Většina těchto technologií předpokládá, že mluvčí při lži budou vykazovat silnější afektivní stavy, např. budou více ve stresu než ti, kdo říkají pravdu, tento přímý vztah však neplatí univerzálně pro všechny mluvčí – i jedinci, kteří říkají pravdu, mohou vykazovat napětí, především ve strachu, že jejich výpověď bude vyhodnocena jako lživá. Naopak mluvčí pronášející lež mohou být klidní a vyrovnaní. Je tedy vhodné zkoumat dopady stresu a lži na řečový projev odděleně (Kirchhübel & Howard, 2011).

Anolli & Ciceri (1997) předkládají model intenční hierarchie (intention hierarchy model) rozlišující dvojí typ situací, za kterých může lež být pronášena, a jejich výzkum ukazuje různé dopady lži na řečový projev mluvčího v oněch situacích. Mluvčí může svou záměrně nepravdivou výpověď směřovat k pasivnímu, pouze naslouchajícímu posluchači, jenž mluvčího nepodezírá ze lži (připravená lež), nebo k posluchači aktivnímu, který klade otázky, lež očekává a přímo se zaměřuje na její případné odhalení, např. při výslechu (nepřipravená lež). V případě připravené lži má mluvčí možnost vyvinout větší kognitivní úsilí a naplánovat přesné komunikační strategie, přičemž není vystaven příliš silnému stresu. Řečové parametry jsou velmi podobné jako u pravdivých promluv, byly pozorovány delší a obsáhlejší repliky, což naznačuje vyvíjení většího úsilí pro formulaci přesvědčivých tvrzení. Naopak při vyslovování lži nepřipravené má mluvčí menší možnost vynakládat mentální úsilí na plánování své výpovědi a je vystaven většímu stresu, což má za následek nápadnější změny v řečovém

projevu: zvýšení  $f_0$  a její variability a častější výskyt dysfluencí, tj. projevy obecně spojené s úzkostí a stresem.

Mluvčí může svůj hlasový projev taktéž měnit zcela záměrně, a to např. při imitaci, kterou lze chápat jako určitou formu maskování hlasu. Imitátor se zaměřuje na určité prominentní jevy v řeči napodobovaného jedince a ty se snaží napodobit. Pomocí analýzy imitace hlasu je možné pozorovat, které aspekty hlasu musí být pozměněny, aby bylo „klamání“ posluchače úspěšné. Lze sledovat, které jevy imitátor mění směrem k cílovému hlasu a které naopak zanechává beze změny, a tudíž umožňují identifikovat, že se jedná o imitátora, nikoliv o napodobovanou osobu. Výzkum Zetterholm (2007) ukázal, že imitátoři poměrně úspěšně napodobují  $f_0$  cílové osoby, stejně jako dialektální intonační vzorce i segmentální jevy. S různou mírou úspěšnosti se pak zhošťují napodobování mluvního tempa. Snaží se také napodobovat mluvní styl typický pro cílového mluvčího a užívat pro něj charakteristické výrazy a hezitační prvky. Schopnost imitátorů napodobovat barvu hlasu cílového mluvčího se různí.

Ukazuje se, že imitátoři se mnohdy dokáží výrazně přiblížit hlasu cílového mluvčího, trénovaný posluchač (např. forenzní fonetik) je však pečlivým poslechem schopen odhalit skutečnost, že se jedná o imitaci (Zetterholm, 2007). U netrénovaných posluchačů se schopnost detekovat imitaci různí. Spolehlivost svědka při rekognici může být negativně ovlivněna jeho věkem, starší mluvčí (nad 60 let) méně často rozpoznají, že se jedná pouze o imitaci hlasu cílového mluvčího. Naproti tomu pohlaví posluchače ani jeho dialekt zřejmě nemají na schopnost detekce imitace signifikantní dopad. Tomu, aby posluchači uvěřili, že slyší projev cílové osoby, nikoliv imitátora, napomáhá hovor na téma, které je s cílovou osobou silněji asociováno (Eriksson et al., 2010).

#### 4. Maskování hlasu

Pokud si je pachatel při činu vědom možnosti, že jeho řečový projev bude nahráván, může se pokusit o maskování svého hlasu, které v budoucnu zkomplikuje jeho identifikaci na základě pořízené nahrávky. Tyto případy nejsou obecně příliš frekventované, Masthoff (1996) uvádí, že v letech 1988-1995 méně než 5 % všech případů trestné činnosti, které byly podrobeny zkoumání na univerzitě v Trieru, obsahovalo maskování hlasu. Frekvence výskytu tohoto fenoménu však závisí na typu trestné činnosti, v některých případech má poměrně značnou pravděpodobnost: nejčastěji se objevuje při únosech, vydírání, sexuálním obtěžování a při falešných telefonických oznámeních, v tomto pořadí (Künzel, 2000). Podle statistik, jež předkládá Masthoff (1996), zaznamenala Německá federální kriminální policie, která se zaměřuje mimo jiné na vydírání a únosy, v letech 1989-1994 výskyt maskování hlasu v 52 % případů, kdy pachatel při telefonátu využil svůj vlastní hlas a mohl předpokládat, že bude nahráván. Pouze v rámci případů vydírání pak poměr využití maskování hlasu dosáhl až 69 %.

Velmi spolehlivým způsobem maskování, který může zásadně zkomplikovat, případně i zcela znemožnit identifikaci mluvčího, je využití elektronického zařízení, tyto případy jsou však jen zřídka; během sedmi let forenzní práce na univerzitě v Trieru nastalo takových případů méně než 1 % (Masthoff, 1996). V poslední době se však objevilo využívání hlasu jiného mluvčího k syntéze připraveného projevu, což identifikaci pachatele na základě hlasu zcela vylučuje (Künzel, 2000).

Většina případů maskování hlasu ve forenzní praxi je méně sofistikovaná. Nejčastějšími způsoby, jak mluvčí mění svůj hlas, jsou zvýšení  $f_0$  (případně změna hlasivkového rejstříku a přechod až do falzetu), snížení  $f_0$  (případně až třepená fonace), šepot, napodobování cizineckého přízvuku a denazalizace pomocí ucpání nosu (Künzel, 2000). Dalšími užívanými metodami jsou např. umístění cizího objektu v ústech, přiložení kapesníku k ústům, omezení pohybu čelisti (např. pomocí přidržování tužky mezi předními zuby), redukce či naopak zvýšení variability  $f_0$ , chrapot či imitace dialektu (Künzel, 2000; Eriksson, 2010).

Künzel (1995; citováno v Masthoff, 1996) a Gfroerer (1994; citováno v Masthoff, 1996) zařazují strategie při maskování hlasu do tří kategorií: (1) změny fonace (do této kategorie spadají změny nastavení hlasivek, jakými jsou změny  $f_0$ , šepot či změny hlasivkového rejstříku), (2) změny jazykového systému (napodobování dialektu či cizineckého přízvuku) a (3) změny řečového projevu (změny tempa, intonace či

artikulačního nastavení). Gfroerer (1994; citováno v Masthoff, 1996) zmiňuje statistickou distribuci forem maskování: v jím analyzovaném vzorku forenzních případů obsahujících maskování hlasu jednu nebo více změn fonace obsahovalo 60 %, jedna nebo více změn jazykového systému se objevilo ve 42 % případů, pouze ve 22 % se vyskytly změny řečového projevu.

Masthoff (1996) uskutečnil výzkum, na jehož základě popsal upřednostňované formy maskování hlasu. Výzkumu se účastnilo 20 univerzitních studentů (13 mužů a 7 žen). Mluvčí volili jeden nebo dva způsoby maskování zároveň, ne však více, převažovaly změny pouze jediného parametru – takový případ nastal v 55 %. Skutečnost, že mluvčí při maskování nemění více parametrů, podle Masthoffa může pramenit z toho, že mluvčí se zároveň s maskováním zaměstnává také potřebou předat informaci, což vyžaduje poměrně značné kognitivní úsilí. Pokud mluvčí maskovali svůj hlas pomocí změny dvou parametrů zároveň, jednalo se vždy o změny v rámci dvou různých kategorií. Ve 35 % případů se mluvčí uchýlovali ke změně fonace, společně se změnou jazykového systému se pak objevila v 15 % případů. 20 % mluvčích zvolilo změnu artikulačního nastavení. Pouze tři z dvaceti způsobů maskování, které mluvčí zvolili, byly efektivní při zastírání vlastností přirozeného řečového projevu daného mluvčího. Masthoff podotýká, že mluvčí, kteří měli při běžném projevu silný regionální přízvuk, se jej při maskování nepokoušeli redukovat. Většina způsobů maskování byla hodnocena jako překvapivě neefektivní samotnými autory, ačkoliv v průběhu nahrávání se mluvčí domnívali, že jimi zvolené způsoby změny hlasu budou mít výraznější dopad.

Nezávisle na způsobu maskování má mluvčí často tendenci míru změny daného parametru postupně snižovat, např. uskutečňuje-li v průběhu času více telefonátů. V některých případech může nakonec maskování i zcela vymizet (Künzel, 2000).

Shinan & Almeida (1986; citováno v Künzel, 2000) zjistili, že na rozdíl od parametrů, které se prostřednictvím (přirozeného, nikoliv elektronického) maskování hlasu dají měnit – jakými jsou i frekvence vokálních formantů, jež jsou obecně využívány jakožto velmi spolehlivé indikátory identity mluvčího – jsou naopak i při maskování velmi stabilní formantové tranzienty, a to zejména ty dlouhé (u diftongů či vokalizovaného /r/), mohou tedy poskytovat idiosynkratické informace o mluvčím.

#### **4.1. Změny fonace**

Jednou z kategorií parametrů, které lze v rámci maskování hlasu modifikovat, jsou prvky související s kmitáním hlasivek při produkování znělých segmentů, tj. s fonací.

Do této kategorie spadá zvyšování či snižování základní frekvence, změny způsobu fonace a šepot. Při zvyšování základní frekvence může dojít až ke změně hlasového rejstříku a mluvčí může přejít do falzetu, při kterém nedochází k úplnému závěru hlasivek a kmitá pouze jejich přední část. Naopak snížením  $f_0$  může mluvčí dosáhnout třepené fonace, při němž jsou hlasivky po většinu času sevřeny a dochází k jejich nepravidelnému rozevírání. Lze využít také napjaté nebo dyšné fonace.

Ve forenzní praxi se nejčastěji objevují způsoby maskování hlasu zahrnující změny základní frekvence. Pro mluvčí jsou tyto změny výhodné, neboť jsou snadno proveditelné, nezpůsobují nesrozumitelnost hlasu, tudíž neomezuji předávání potřebných informací, a zároveň jsou dostatečně efektivní pro skrytí identity mluvčího (Künzel, 2000).

Podle studie Masthoffa (1996) při maskování hlasu zvyšování  $f_0$  (včetně falzetu) využívali pouze muži, zatímco jejího snižování (včetně třepené fonace) se dopouštěly výhradně ženy. Kromě tohoto výsledku nebyl v rámci jeho výzkumu nalezen vztah mezi zvolenými způsoby maskování a vlastnostmi modálního hlasu mluvčích. Künzel (2000) však na základě svého experimentu, v němž participovalo 100 mluvčích (50 mužů a 50 žen) dospěl k závěru, že mluvčí s vyšší přirozenou  $f_0$  mají tendenci k jejímu dalšímu zvyšování a mluvčí s nižší přirozenou  $f_0$  ji naopak ještě více snižují, a to až k třepené fonaci. Ženy se v rámci jeho experimentu při maskování hlasu uchýlovaly k méně dramatickým změnám své  $f_0$ , muži ji obecně měnili výrazněji, častěji se uchýlovali k falzetu. Při snižování základní frekvence však výraznější rozdíly vykazovaly ženy (3,5 půltónu, zatímco muži se při snižování  $f_0$  dostávali níže v průměru o 2,5 půltónu), Künzel (2000) navrhuje vysvětlení obecným rozdílem  $f_0$  mužů a žen: muži mají nižší přirozenou  $f_0$  a tudíž více prostoru pro její zvyšování, u žen je tomu naopak. V Künzelově studii, na rozdíl od studie Masthoffovy, muži častěji volili snižování  $f_0$ , zatímco ženy preferovaly její zvyšování.

McGehee (1937; citováno v Eriksson, 2010) experimentálně testovala úspěšnost rekonstrukce při změně  $f_0$  mluvčího. Správnost identifikace byla snížena z 80 % (v případě, že respondenti slyšeli mluvčího hovořit pouze nemaskovaně) na 63 % (pokud respondenti dříve slyšeli mluvčího hovořit se změněnou  $f_0$ ). Podobný experiment provedli Wagner & Köster (1999; citováno v Eriksson, 2010), kteří pouštěli respondentům nahrávky jim známých mluvčích hovořících modálně a falzetem. Respondenti měli následně určit mluvčího. Pokud mluvčí nemaskovali svůj hlas,

dosahovala úspěšnost identifikace 97 %, zatímco v případě, že mluvčí hovořili falzetem, byla správnost identifikace snížena na pouhých 4 %.

Künzel (2000) na základě výsledků svého výzkumu uvádí, že pokud je maskování dosahováno pomocí snižování průměrné  $f_0$ , je možné přirozenou základní frekvenci mluvčího do jisté míry odhadnout na základě maskovaných hodnot. Pokud však mluvčí při maskování  $f_0$  zvyšuje, nejsou k dispozici silné inference z přirozené  $f_0$ , a to nezávisle na tom, zda dojde ke změně hlasového rejstříku, či nikoliv.

Při extrémních změnách  $f_0$  se mění také vokální formanty, ačkoliv obvykle jsou na sobě  $f_0$  a formanty více méně nezávislé. Výrazné snížení vede k poklesu  $F_1$ ,  $F_2$  a  $F_3$ , naopak je tomu při zvýšení  $f_0$ . Tato skutečnost je zřejmě zapříčiněna prodloužením vokálního traktu při snížení hrtanu (při snižování  $f_0$ ) a naopak jeho zkrácením při zvýšení hrtanu (Shinan & Almeida; citováno v Künzel, 2000). Při zvyšování či snižování  $f_0$  dochází také k poklesu artikulačního tempa a nárůstu množství a trvání pauz, což je pravděpodobně dáno tím, že se mluvčí soustředí na neobvyklé nastavení hlasivek, a nemůže tudíž věnovat plnou pozornost obsahu svého sdělení. Často se také objevuje změna hlasitosti – její nárůst při zvýšení  $f_0$  a pokles při snížení  $f_0$  (Künzel, 2000).

Běžně se v řeči vyskytuje třepená fonace, zejména na konci nádechových úseků, většinou ovšem není dominantním způsobem fonace. Pokud se k ní mluvčí uchýlí, lze tuto skutečnost snadno detekovat poslechem. Vyznačuje se nepravidelnými vibracemi hlasivek, nízkou  $f_0$  a nízkou amplitudou kmitání, tj. není příliš hlasitá. Třepená fonace je snadno udržitelná po poměrně dlouhou dobu. V řečovém projevu se obvykle objevuje pouze v kratších úsecích, pokud je v ní ale mluvčí konzistentní, lze ji využít k maskování hlasu (Eriksson, 2010). Hirson & Duckworth (1995) testovali úspěšnost identifikace mluvčího pomocí poslechu, nahrávky mluvčích, kteří hovořili modální a třepenou fonací, byly předloženy respondentům, mezi nimiž byli jak laici, tak i školení fonetici. Při užití třepené fonace mluvčím byla úspěšnost rekognice laiky snížena z 93 % na 51 %, úspěšnost školených fonetiků se ve stejném případě snížila z 99 % na 73 %. Výsledky tedy potvrdily, že foneticky trénovaní posluchači vykazují při rekognici lepší výsledky, a to i v případě, že mluvčí maskuje svůj hlas.

Mluvčí může ve svém řečovém projevu zcela eliminovat základní frekvenci, pakliže hovoří šepem. Dopad šepotu na úspěšnost rekognice zkoumali Orchard & Yarmey (1995). Respondentům byl nejprve prezentován hlas figuranta, který hovořil běžným hlasem nebo šepem. Později jej měli respondenti identifikovat mezi dalšími pěti



figuranty. Respondenti, kteří prve slyšeli figurantův běžný hlas, prováděli rekognici mezi nahrávkami běžných hlasů všech figurantů, zatímco posluchači, kteří původně slyšeli jeho šepot, vybírali jeho hlas mezi šeptanými nebo běžně pronášenými projevy. Nejúspěšněji si při rekognici počínali respondenti, kteří prve i při rekognici slyšeli figurantův běžný projev, horších výsledků dosáhli posluchači, kteří v obou případech slyšeli šepot, a nejnižší míru úspěšnosti vykazali ti, kteří nejprve slyšeli figurantův šepot a následně prováděli rekognici mezi nahrávkami běžných projevů.

#### **4.2. Napodobování dialektu a cizineckého přízvuku**

Jak již bylo zmíněno v oddíle 2.6, při profilování mluvčího i při porovnávání sporné a srovnávací nahrávky je užitečným faktorem dialekt. Při rekognici je zásadní, aby dialekt podezřelého byl shodný s dialektem zúčastněných figurantů, či mu byl alespoň velice blízký. Dialekt může být natolik silnou součástí identity mluvčího, že jím lze zastřít další vlastnosti projevu mluvčího (Eriksson, 2010). Sjöström et al. (2006; citováno v Eriksson, 2010) provedli experiment s mluvčími ovládajícími dva švédské dialekty, respondenti následně prováděli jejich identifikaci mezi čtyřmi figuranty. Mluvčí byl identifikován se stejnou mírou úspěšnosti, hovořil-li na sporné i srovnávací nahrávce tímž dialektem. Pokud jej však posluchači nejprve slyšeli mluvit jedním dialektem a při rekognici následně hovořil druhým z nich, úspěšnost identifikace se propadla na úroveň náhody. Autoři experimentu považují za možný důvod skutečnost, že při posuzování identity mluvčího je dialekt velice silným faktorem a má vyšší prioritu než ostatní jevy. Pokud je tomu tak, Eriksson (2010) se domnívá, že imitace dialektu, která úspěšně zakryje identitu mluvčího, avšak nezní autenticky, bude mít slabší vliv na úspěšnost rekognice.

Kromě dialektu lze v rámci maskování hlasu napodobit také cizinecký přízvuk. Neuhauser (2008) pozorovala německé mluvčí imitující francouzský přízvuk, pro analýzu zvolila několik charakteristických parametrů. Mluvčí byli při jejich imitaci většinou nekonzistentní, úspěšnost jejich napodobení se různila, v některých případech si mluvčí ony jevy vůbec neuvědomovali. Většinou docházelo ke změnám artikulace, polovina mluvčích modifikovala také základní frekvenci svého hlasu. Neuhauser & Simpson (2007) sledovali schopnost rodilých mluvčích němčiny identifikovat autentický a imitovaný francouzský přízvuk, tj. určit, zda je autorem předložené nahrávky rodilý mluvčí francouzštiny hovořící německy, či rodilý mluvčí němčiny napodobující francouzský přízvuk. Rodilí mluvčí němčiny se lišili ve schopnosti

napodobit cizinecký přízvuk tak, aby byl posluchači vyhodnocen jako autentický. Překvapivým výsledkem byly značné rozdíly ve schopnosti posluchačů úspěšně identifikovat mluvčího hovořícího s autentickým francouzským přízvukem. Někteří nerodilí mluvčí němčiny tedy hovořili s méně přesvědčivým cizineckým přízvukem než mluvčí rodilí, kteří přízvuk pouze imitovali.

#### **4.3. Cizí objekt v ústech**

Mimo způsoby maskování využívající pouze změn nastavení vokálního traktu mluvčího se lze setkat také s maskováním hlasu pomocí cizího objektu umístěného v ústech, případně před ústy, před nimiž si pachatel může např. přidržovat kapesník (Künzel, 2000).

Figueiredo & Britto (1996) popsali dopad přidržování tužky mezi předními zuby na řečový projev mluvčího. Tímto způsobem je omezena artikulace hlásek, přičemž různé segmenty jsou ovlivňovány do různé míry, nelze tedy jejich změny popsat plošně. Autoři se zaměřili zejména na artikulaci vokálů, při tomto způsobu maskování je percepčně nejvýraznější snížení vysokých vokálů způsobené znehybněním spodní čelisti. Rty jsou povětšinou otevřené a je omezen jejich pohyb, je komplikováno zaokrouhlování vokálů. Vokalický prostor je zmenšen a vokály jsou zpravidla do jisté míry kvalitativně redukovány. Je omezen pohyb jazyka, jemuž je kvůli cizímu tělesu v ústech zabráněno v pohybu do anteriorní oblasti. Autoři uvádí, že v brazilské portugalské je dopad maskování na vokály zásadní, neboť výslovnost jednotlivých vokálů je charakteristická pro některé regionální dialekty a maskování těchto jevů může zkomplikovat profilování a identifikaci mluvčího. Kromě změn akustických charakteristik vokálů existuje také předpoklad dopadu tohoto typu maskování na konsonanty, zejména posteriorizace /s/ v [ʃ], což může také znesnadnit identifikaci, neboť posteriorizace tohoto segmentu je taktéž charakteristická pro některé dialekty brazilské portugalské.

Jakákoliv maskovací strategie, která modifikuje supralaryngální charakteristiky mluvčího, nikoliv pouze jeho fonaci, má za následek značnou komplikaci identifikace, neboť efekt takového maskování je obtížně předvídatelný, zejména pro komplexnost interakce omezení způsobeného umístěním cizího předmětu v ústech a přirozeného artikulačního nastavení mluvčího. Při maskování, které zahrnuje změny fonace, je většina dialektálních a segmentálních jevů zachována, což umožňuje poměrně efektivně odhadnout dialekt při profilování mluvčího. Při maskování hlasu pomocí cizího objektu

v ústech je odhadnutí podoby běžného projevu mluvčího obtížnější, jednou z možností je identifikace změn, které podstupuje každý z typů segmentů. Lze také nahrát projev podezřelého s tužkou svíranou mezi předními zuby a porovnat takový projev se spornou nahrávkou (Figueiredo & Britto, 1996).

#### **4.4. Elektronické maskování hlasu**

Elektronické maskování hlasu ve forenzní praxi není příliš častým jevem, Gfroerer (1994; citováno v Clark & Foulkes, 2007) odhaduje jeho výskyt na cca 10 % všech případů obsahujících maskování hlasu, zatímco Masthoff (1996) jej nalézá v méně než 1 % případů. Přesto je však vhodné se tímto fenoménem zabývat, Clark & Foulkes (2007) předpokládají, že jeho četnost bude postupně vzrůstat, neboť zařízení uzpůsobená k elektronickému maskování hlasu jsou čím dál dostupnější. Tento způsob maskování hlasu lze využívat také např. k ochraně identity svědků v soudních procesech, pro takové účely je zásadní, aby bylo maskování dostatečně efektivní pro znemožnění identifikace mluvčího, a to i posluchačem, kterému je jeho hlas znám, a zároveň by nemělo mít negativní dopad na srozumitelnost projevu mluvčího.

Clark & Foulkes (2007) se ve své studii zaměřili na dopad elektronického maskování hlasu pomocí modifikace základní frekvence na identifikaci mluvčího. Autoři se rozhodli pro modifikaci  $f_0$ , neboť je nejčastěji zasaženým parametrem při využití běžných přirozených způsobů maskování (viz oddíl 4.1). Základní frekvence je také snáze modifikovatelná než jiné parametry, jakými jsou např. segmentální jevy či barva hlasu. Respondentům byly při experimentu předloženy nahrávky projevu čtyř mluvčích, s jejichž hlasy se měli obeznámit, aby je později byli schopni identifikovat. Následně prováděli rekognici, a to mezi přirozenými a uměle modifikovanými nahrávkami. Úspěšnost identifikace byla s maskováním snížena, přičemž s větší změnou  $f_0$  úspěšnost klesala, u většiny posluchačů však byla zachována míra úspěšnosti identifikace nad úroveň náhody, a to i v případě nejvýraznější změny  $f_0$ , tj. posunu o 8 půltónů. Míra úspěšnosti identifikace se při běžném projevu i při aplikaci elektronického maskování různila u jednotlivých mluvčích. Efektivita maskování nebyla univerzální, jeho vliv na identifikovatelnost jednotlivých mluvčích se lišil.

## **5. Strategie maskování hlasu u českých mluvčích**

Cílem této práce je popis metod, které využívají čeští mluvčí při maskování svého hlasu, jsou-li instruováni, aby jej maskovali zcela libovolným způsobem, tak, aby nebyli podle hlasu identifikovatelní. Budou předloženy jednotlivé způsoby, jež mluvčí využívají k maskování svého projevu, a jejich akustická analýza. Efektivita maskování mluvčích, kteří zvolili komplexnější strategie, bude vyhodnocena pomocí percepčních testů, v nichž budou respondenti hodnotit zaprvé nápadnost maskování, tj. zda je zřejmé, že se jedná o maskovaný projev, či nikoliv, a zadruhé identifikovatelnost mluvčího na základě hlasu.

Cílem popisu způsobů maskování hlasu není jednoznačná forenzní identifikace mluvčího. Tyto jednotlivé poznatky mohou napomoci zúžení rozsahu možných podezřelých a podpořit identifikaci mluvčího v kombinaci s dalšími nejen fonetickými metodami.

## **6. Metoda**

### **6.1. Materiál**

Byly pořízeny nahrávky celkem 80 mužů, rodilých mluvčích obecné češtiny ve věku 18 – 40 let. Tyto nahrávky jsou součástí referenční databáze pro forenzní účely. Získávání řečového materiálu probíhalo při nahrávacích sezeních, při nichž bylo pořizováno několik typů projevů mluvčích, trvajících asi 45 minut. Nahrávání bylo uskutečňováno v tichém prostředí za použití diktafonu Edirol R09 a byl pořizován signál o vzorkovací frekvenci 48 kHz.

Součástí nahrávacího sezení byla úloha, při níž mluvčí četli dva texty, jež byly následně využity v této studii. Přednes každého z textů trval asi jednu minutu. První z textů (text A) mluvčí předčítali svým běžným hlasem, druhý z nich (text B) pak hlasem maskovaným libovolným způsobem, přičemž byli instruováni, aby změnili svůj hlas tak, aby nebyli rozpoznatelní na základě pořízené nahrávky. Texty A a B se lišily, obsahovaly však některé shodné fráze.

### **6.2. Procedura**

Nejprve byla uskutečněna poslechová analýza nahrávek maskovaných i nemaskovaných projevů všech mluvčích za účelem celkového zmapování strategií užívaných při maskování hlasu. Na základě této analýzy bylo následně vybráno 15 mluvčích, u nichž byla provedena také akustická analýza a jejichž projevy byly vyhodnocovány pomocí percepčních testů.

#### **6.2.1. Poslechová analýza**

Byly porovnány nahrávky maskovaných a nemaskovaných projevů všech mluvčích a následně zaznamenány strategie, které mluvčí volili při maskování svého hlasu: změna základní frekvence (snížení, zvýšení, beze změny), způsobu fonace (tlačená, dyšná, třepená, šepot, modální) a tempa (pomalejší, rychlejší, stejné), případně další způsoby změny řečového projevu (nazalizace, změna artikulace, změna melodického průběhu atd.).

Na základě poslechu a porovnání nemaskovaného projevu mluvčího s maskovaným byly hodnoceny obtížnost rozpoznání (snadné, střední, obtížné) a dojem, kterým maskovaný projev působil (přirozený, nepřirozený), tj. zda bylo při poslechu maskovaného projevu zřejmé, že mluvčí záměrně mění svůj hlas, či nikoliv. (Např. změny artikulace působily ve většině případů přirozeně, zatímco v případě extrémních

změn fonace – výrazném zvýšení, případně až falzetu, či naopak snížení vedoucím až k třepené fonaci – se jednalo o nápadné modifikace.)

Tyto ukazatele byly následně hodnoceny respondenty při percepčních testech pro 15 vybraných mluvčích (viz oddíl 6.2.3).

### **6.2.2. Akustická analýza**

Na základě výsledků poslechové analýzy byl vytvořen výběr 15 mluvčích, jejichž maskovaný projev byl obtížně rozpoznatelný nebo rozpoznatelný středně a působící přirozeně.

U těchto mluvčích byla následně provedena akustická analýza kvantifikující rozdíly mezi přirozeným a maskovaným řečovým projevem.

V programu Praat byly vyznačeny hranice segmentů v jednotlivých nahrávkách, nejprve automaticky pomocí nástroje Prague Labeller (Pollák et al., 2007), následně byla provedena manuální korekce.

Pomocí autokorelační funkce v Praatu (s intervalem 10 ms a rozsahem extrakce 60-350 Hz) byla pro jednotlivé řečové vzorky vypočítána střední základní frekvence a její směrodatná odchylka.

Pro jednotlivé vokalické kvality v rámci nahrávek byly změřeny hodnoty formantů F1, F2 a F3 v rámci rozsahu 0-5000 Hz, a to v několika ekvidistantních bodech v průběhu daného vokálu v intervalu 10 ms. Byly analyzovány vokály /ɪ/, /e/, /a/, /o/, /u/, /i:/, /a:/ a /u:/; do výpočtu nebyly zahrnuty diftongy a také vokály /e:/ a /o:/, které se v textu, jež mluvčí předčítali maskovaným hlasem, neobjevily.

Následně byl pro analyzované vokály vyslovované jednotlivými mluvčími proveden t-test pro zjištění signifikance rozdílů v jejich realizaci při maskovaných a nemaskovaných projevech.

Pro každého mluvčího bylo vypočteno mluvní a artikulační tempo při maskovaném i nemaskovaném projevu.

Získané hodnoty byly následně porovnávány pro maskované a nemaskované projevy mluvčích.

### **6.2.3. Percepční testy**

Z nahrávek projevů 15 vybraných mluvčích byly extrahovány čtyři fráze, jež se objevovaly v obou textech. Z těchto frází byly sestaveny dva percepční testy. Jednotlivé položky byly kvalitativně normalizovány kvůli nestejným podmínkám nahrávání

jednotlivých mluvčích, které zapříčinily drobné slyšitelné rozdíly v kvalitě nahrávek a mohly by ovlivnit výsledky percepčních testů.

Skupinu respondentů tvořilo 15 univerzitních studentů filologických oborů mezi 20 a 30 lety.

V obou textech se nacházely čtyři shodné fráze, každá z testovacích položek byla sestavena ze dvou frází čtených maskovaným či nemaskovaným hlasem. Celkem tedy bylo vytvořeno 60 testovacích položek – od každého mluvčího po dvou maskovaných a nemaskovaných položkách.

Posluchači prováděli percepční testy doma pomocí programu Praat, přičemž byli instruováni, aby využili sluchátek a před zahájením testu se ujistili o vhodném zvukovém nastavení svého počítače.

V rámci prvního testu byly posluchačům v náhodném pořadí přehrávány maskované i nemaskované položky a respondenti hodnotili na pětistupňové Likertově škále, zda jim daný projev připadá maskovaný, či ne; volili odpověď z variant „určitě ne“ (-2), „spíše ne“ (-1), „těžko říct“ (0), „spíše ano“ (+1) a „určitě ano“ (+2). Každá z položek byla posluchači přehrána pouze jednou, mezi jednotlivými položkami vždy zazněla krátká desenzitační pasáž. Tímto způsobem bylo testováno, zda daný způsob maskování hlasu působí přirozeně, nebo ne. Hodnoceny byly i nemaskované položky kvůli následnému porovnávání s výsledky pro maskované položky jednotlivých mluvčích. Na základě tohoto testu bylo získáno 60 odpovědí od každého posluchače, neboť hodnoceny byly všechny položky.

V rámci druhého testu posluchači prováděli rekognici. Byly jim postupně předkládány maskované položky od všech mluvčích, které měli spojit s jejich běžnými hlasy (zachycenými na nemaskovaných položkách), jež volili ze tří možností; jako figuranti byli vždy využiti dva další mluvčí, vybraní na základě podobnosti přirozené f0 a barvy hlasu. Test obsahoval 15 oddílů, posluchači prováděli rekognici pro každého mluvčího zvlášť a každý mluvčí byl v testu reprezentován pouze jednou. Srovnávacím (nemaskovaným) nahrávkám byla náhodně přiřazena čísla 1-3 a posluchači označovali číslo srovnávací položky, o které se domnívali, že jejím autorem je týž mluvčí, který hovoří na položce sporné. Nebyla zahrnuta možnost nepřítomnosti srovnávací (nemaskované) nahrávky mluvčího zachyceného na sporné (maskované) nahrávce. Respondentům byl umožněn libovolný počet přehrávání maskovaných i nemaskovaných položek. Celkově bylo získáno 15 odpovědí od každého posluchače. Tento test ověřil míru obtížnosti identifikace mluvčího na základě jeho maskovaného projevu.

Posluchači byli instruováni, aby neměnili pořadí percepčních testů, kvůli možnosti zapamatování maskovaných projevů jednotlivých mluvčích předložených při rekognici ve druhém testu.



## 7. Výsledky

### 7.1. Poslechová analýza

Nahrávky všech 80 mluvčích byly nejprve podrobeny poslechové analýze, při níž byly zaznamenány jednotlivé strategie využití k maskování hlasu. Hodnocenými parametry byly (1) základní frekvence, (2) fonace, (3) tempo a (4) případné další strategie. Analýza spočívala v opakovaném poslechu nahrávek a pečlivém srovnávání maskovaných a nemaskovaných projevů mluvčích.

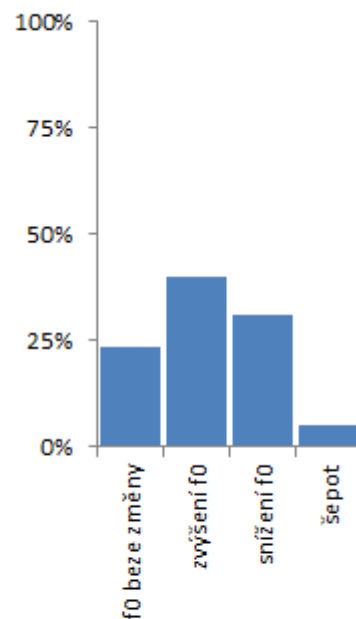
Změnu střední základní frekvence při maskování hlasu zvolila většina mluvčích. Nejčastěji se jednalo o její zvýšení, pro nějž se rozhodlo 32 mluvčích. Svou střední základní frekvenci naopak snížilo 25 mluvčích, z toho 4 přešli do třepené fonace. Bez postřehnutelné změny zůstala střední  $f_0$  u 19 mluvčích.

Zbývajících 4 mluvčích hovořili při maskování hlasu šepem, základní frekvence u nich tudíž byla zcela eliminována.

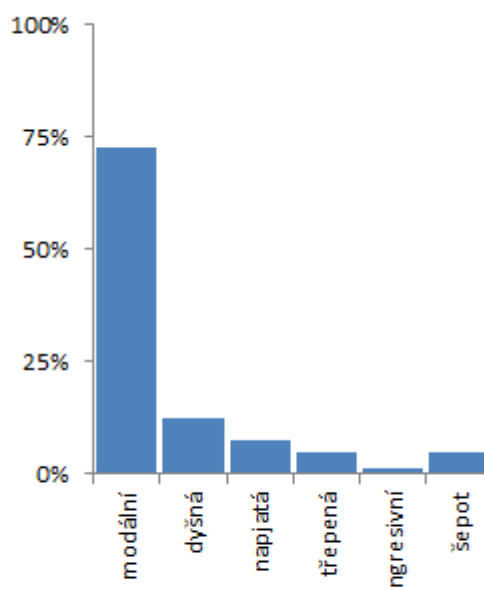
U 17 mluvčích se objevila pouze změna tohoto parametru. V 6 případech se jednalo o snížení střední  $f_0$ , naopak v 11 případech mluvčích svou střední  $f_0$  zvýšili.

Způsoby modifikace  $f_0$  zvolené mluvčími a jejich poměr znázorňuje graf 7-1.

Pro změnu fonace se při maskování rozhodla menšina mluvčích, 58 zůstalo u své modální fonace. U 10 mluvčích se objevila dyšná a u 6 mluvčích napjatá fonace. 4 mluvčích maskovali pomocí třepené fonace a 4 mluvčích se rozhodli pro šepot (viz předchozí odstavec). V 1 případě se objevila ingresivní fonace. Fonační nastavení jednoho z mluvčích kolísalo mezi fonací třepenou a modální. U žádného z mluvčích nebyl zaznamenán falzet.



**Graf 7-1:** změny  $f_0$  v rámci maskování hlasu

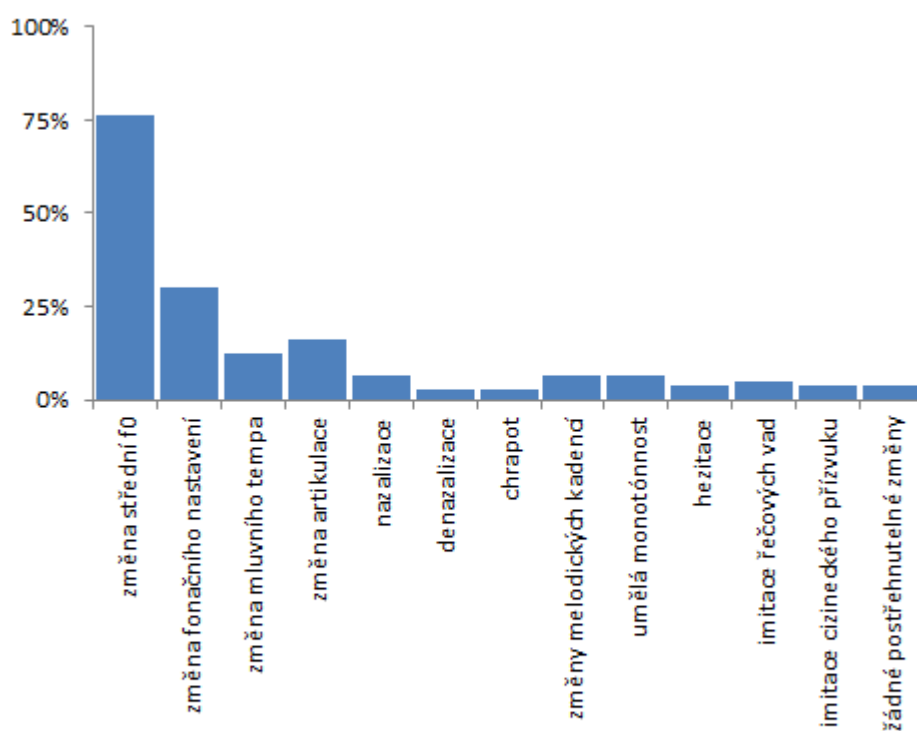


**Graf 7-2:** způsoby fonace využití při maskování hlasu

Pouhá změna fonace (a žádného jiného parametru, mimo  $f_0$  v případě třepené fonace, kdy nevyhnutelně došlo k jejímu snížení) byla zaznamenána u 3 mluvčích: 1 z nich zvolil třepenou, 1 napjatou a 1 ingresivní fonaci. Jednotlivé způsoby fonace využití mluvčími a jejich četnost zobrazuje graf 7-2.

Modifikaci svého mluvního tempa uskutečnilo ještě méně jedinců než změnu fonace, tempo bez postřehnutelného rozdílu oproti běžnému projevu bylo při maskování zachováno u 70 mluvčích. Nižší tempo se objevilo u 8 mluvčích, pouze 2 mluvčí se rozhodli pro jeho zvýšení. Žádný z mluvčích neuskutečnil změnu mluvního tempa jakožto jediného parametru.

Ve 13 případech se mluvčí rozhodli pro změnu artikulace, z toho 4 zvolili faryngalizaci, 1 přednější, 1 zavřenější a 1 otevřenější artikulaci vokálů. Dalšími strategiemi, které mluvčí při maskování hlasu využili, byly nazalizace (5 mluvčích), denazalizace pomocí ucpání nosu (2 mluvčí), chrapot (2 mluvčí), změny melodických kadencí coby realizací melodémů ukončujících klesavých (5 mluvčích), redukce variability  $f_0$ , tj. umělá monotónnost (5 mluvčích), nápadná produkce pauz a hezitačních prvků (působící jako vědomá změna oproti nemaskovanému čtenému projevu; 3 mluvčí), napodobování řečových vad (4 mluvčí) a imitace cizineckého



**Graf 7-3:** souhrn četnosti všech modifikací využitých mluvčími při maskování hlasu

přízvuku (3 mluvčí). Ve zkoumaném vzorku se neobjevily imitace dialektu ani hovor s cizím předmětem v ústech. U 3 mluvčích se neobjevily postřehnutelné změny žádného parametru.

Souhrn četností všech modifikací využitých mluvčími při maskování hlasu, které byly zaznamenány při poslechové analýze, představuje graf 7-3.

Co se týče hodnocení přirozenosti dojmu maskovaného projevu mluvčích a míry obtížnosti identifikace mluvčího na základě jeho poslechu, u většiny mluvčích bylo rozpoznání hodnoceno jako snadné, jejich počet dosáhl 49, z toho projev 29 z nich působil spíše nepřirozeným dojmem, zatímco 20 mluvčích docílilo projevu působícího přirozeně.

Většina mluvčích, jejichž maskovaný hlas působil nepřirozeně (NP) a zároveň byli na jeho základě snadno identifikovatelní (S), zvolila změnu střední  $f_0$ ; 13 mluvčích se rozhodlo pro její zvýšení, 6 naopak pro snížení (u 10 mluvčích v této kategorii nedoznala střední  $f_0$  výrazných změn). 8 z těchto mluvčích provedlo změnu způsobu fonace, ve 3 případech šlo o fonaci dyšnou, ve 3 o napjatou a ve 2 o třepenou, která ale v jednom případě nebyla přítomna stabilně a mluvčí přecházel také k fonaci modální. 4 mluvčí snížili své mluvní tempo, 1 mluvčí se naopak rozhodl jej zvýšit.

U mluvčích, kteří byli navzdory maskování hlasu snadno rozeznatelní (S), a jejich maskovaný projev působil přirozeně (P), se objevovalo zvýšení  $f_0$  v menší míře, z 20 mluvčích se pro ni rozhodli pouze 4. Polovina mluvčích v této kategorii, tj. 10, svou střední  $f_0$  naopak snížila. U zbývajících 6 zůstala střední  $f_0$  bez znatelných změn. Většina těchto mluvčích při maskování zůstala u své modální fonace, pouze 1 mluvčí hovořil dyšně. Mluvní tempo snížili 2 mluvčí, 1 jej naopak zvýšil.

Z 16 mluvčích, jejichž identifikovatelnost na základě maskovaného hlasu byla hodnocena jako střední (Stř), a jejichž maskovaný projev působil nepřirozeně (NP), došlo u poloviny, tj. u 8 mluvčích, ke zvýšení střední  $f_0$ , naopak její snížení vykázali 4 mluvčí. 2 mluvčí svou střední  $f_0$  ponechali bez výrazné změny. 2 mluvčí v rámci maskování hlasu hovořili šeptem. 8 mluvčích při maskování zůstalo u své modální fonace, 2 mluvčí využili fonaci napjatou a 2 dyšnou. V 1 případě se objevila třepená a v 1 případě ingresivní fonace. Ke snížení mluvního tempa došlo u 2 mluvčích, ostatní hovořili tempem, které se výrazně nelišilo od tempa jejich nemaskovaného projevu.

Mluvčích, jejichž rozpoznatelnost byla hodnocena jako střední (Stř), a jejichž maskovaný projev působil přirozeně (P), se v analyzovaném vzorku objevilo 5. Všichni tito mluvčí zvýšili svou střední  $f_0$ . 3 mluvčí využili dyšné fonace, zbývajících 2 zůstali u

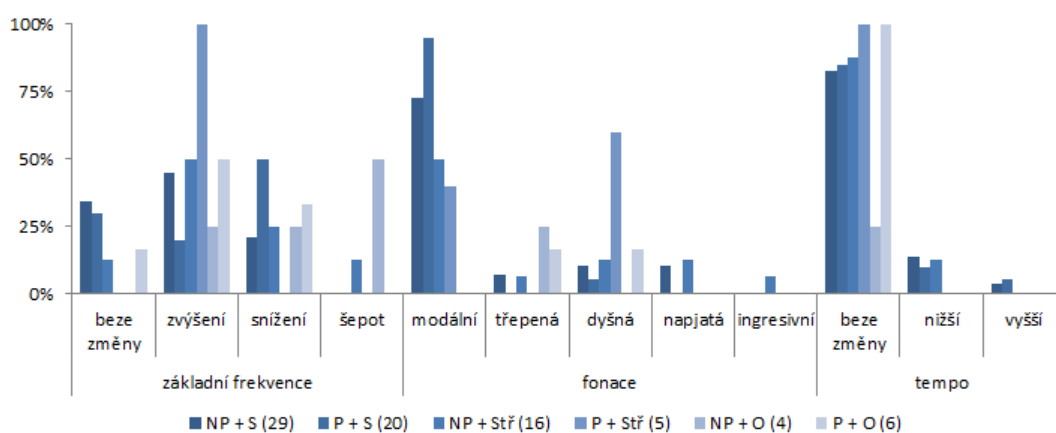
modálního fonačního nastavení. Všichni mluvčí hovořili stejným tempem jako při svém nemaskovaném projevu. 1 z mluvčích v této kategorii změnil svou artikulaci, 1 zařadil do svého projevu hezitační prvky a modifikoval melodii své řeči.

Jako obtížně identifikovatelní na základě maskovaného hlasu (O) byli vyhodnoceni 4 mluvčí, jejichž maskovaný projev zároveň působil nepřírozeně (NP). 2 z těchto mluvčích při maskování šeptali, 1 zvýšil svou střední f0 a 1 ji naopak snížil, a to do té míry, že přešel k třepené fonaci. U žádného z těchto mluvčích nebyla pozorována změna mluvního tempa. 2 z těchto mluvčích změnil svou artikulaci, 1 z nich navíc provedl nazalizaci a druhý hovořil uměle monotónně.

kategorie:		NP + S	P + S	NP + Stř	P + Stř	NP + O	P + O
počet mluvčích:		29	20	16	5	4	6
základní frekvence	beze změny	10	6	2	0	0	1
	zvýšení	13	4	8	5	1	3
	snížení	6	10	4	0	1	2
	šept	0	0	2	0	2	0
fonace	modální	21	19	8	2	0	0
	třepená	2	0	1	0	1	1
	dyšná	3	1	2	3	0	1
	napjatá	3	0	2	0	0	0
	ingresivní	0	0	1	0	0	0
tempo	beze změny	24	17	14	5	4	6
	nižší	4	2	2	0	0	0
	vyšší	1	1	0	0	0	0

**Tabulka 7-1:** četnosti změn f0, fonace a tempa v kategoriích stanovených při poslechové analýze

Jako obtížně identifikovatelní mluvčí (O) s přirozeně působícím maskovaným projevem (P) bylo označeno 6 jedinců. U 2 z nich došlo ke snížení střední f0, přičemž 1 využil třepené fonace, u 3 mluvčích byla f0 naopak zvýšena. U 1 mluvčího zůstala beze změny, tento mluvčí se uchýlil k dyšné fonaci. Mluvní tempo žádného mluvčího se výrazně nelišilo od mluvního tempa při běžném čteném projevu. 5 z těchto mluvčích provedlo modifikace své artikulace, v jednom případě se jednalo o imitaci cizineckého přízvuku, ve dvou pak o imitaci vad řeči (lambdacismus, sigmatismus, rotacismus). Četnosti změn f0, fonace a tempa v jednotlivých kategoriích (NP+S, P+S, NP+Stř, P+Stř, NP+O, P+O) zobrazuje tabulka 7-1, poměry užití pak znázorňuje graf 7-4.



**Graf 7-4:** poměry modifikací parametrů při maskování hlasu v kategoriích stanovených při poslechové analýze

Nahrávky mluvčích spadajících do posledních tří ze jmenovaných kategorií – tj. (1) maskovaný projev působící přirozeně + střední identifikovatelnost mluvčího na základě jeho maskovaného hlasu (P+Stř), (2) nepřirozeně znějící maskovaný projev + obtížná identifikovatelnost (NP+O) a (3) přirozeně znějící maskovaný projev + obtížná identifikovatelnost (P+O) – byly podrobeny akustické analýze a využity při percepčních testech.

## 7.2. Akustická analýza

### 7.2.1. Základní frekvence

U většiny mluvčích při maskování hlasu základní frekvence vzrostla, nejvýraznější rozdíl byl zaznamenán u mluvčího MAKN, jehož střední  $f_0$  byla při maskování zvýšena o 161 Hz (z původních 139 na 301 Hz, tj. o 13,4 půltónu). Nárůst  $f_0$  při maskovaném projevu v řádech desítek Hz vykázali také mluvčí JELS (59 Hz, 5,2 půltónu), VEDR (54 Hz, 6,8 půltónu), SJAN (49 Hz, 6,5 půltónu), OLET (35 Hz, 4,7 půltónu), ZALT (35 Hz, 4,5 půltónu), MAND (28 Hz, 3,3 půltónu) a KUCK (17 Hz, 1,8 půltónu).

U čtyř mluvčích došlo naopak ke snížení  $f_0$  v řádech desítek Hz, jednalo se o mluvčí VLAJ (pokles o 34 Hz ze 133 na 99 Hz, tedy o 5,1 půltónu), KAND (29 Hz, 5 půltónů), ROLK (17 Hz, 2,1 půltónu) a ZENT (14 Hz, 2 půltóny).

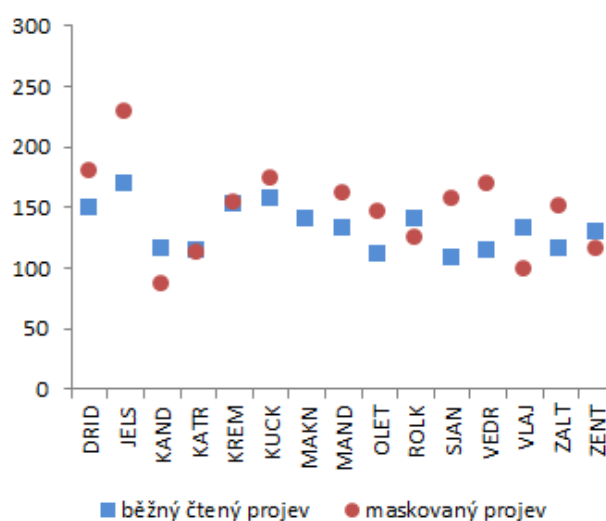
Hodnoty středních  $f_0$  mluvčích při maskovaném a nemaskovaném projevu předkládá tabulka 7-2, jejich porovnání pak znázorňuje graf 7-5.

Variabilitu své střední  $f_0$  výrazně zvýšil mluvčí KUCK, v jehož případě byla směrodatná odchylka  $f_0$  při nemaskovaném projevu rovna 26 Hz, zatímco při projevu maskovaném dosáhla 71 Hz, vzrostla tedy o 45 Hz. Zvýšení variability  $f_0$  bylo zaznamenáno také u mluvčích VEDR (nárůst směrodatné odchylky o 19 Hz) a MAND (12 Hz). Naopak ke snížení variability střední základní frekvence došlo u mluvčího VLAJ, v jehož případě klesla směrodatná odchylka o 17 Hz.

Hodnoty směrodatných odchylek jednotlivých položek jsou vyčísleny v tabulce 7-3, porovnání jejich velikostí představuje graf 7-6.

mluvčí	běžný projev	maskovaný projev	rozdíl
DRID	249 Hz	180 Hz	31 Hz
JELS	169 Hz	228 Hz	59 Hz
KAND	115 Hz	86 Hz	-29 Hz
KATR	113 Hz	112 Hz	-1 Hz
KREM	152 Hz	153 Hz	1 Hz
KUCK	157 Hz	174 Hz	17 Hz
MAKN	139 Hz	301 Hz	161 Hz
MAND	132 Hz	160 Hz	28 Hz
OLET	111 Hz	146 Hz	35 Hz
ROLK	140 Hz	124 Hz	-17 Hz
SJAN	107 Hz	156 Hz	49 Hz
VEDR	114 Hz	169 Hz	54 Hz
VLAJ	133 Hz	99 Hz	-34 Hz
ZALT	116 Hz	150 Hz	34 Hz
ZENT	129 Hz	115 Hz	-14 Hz

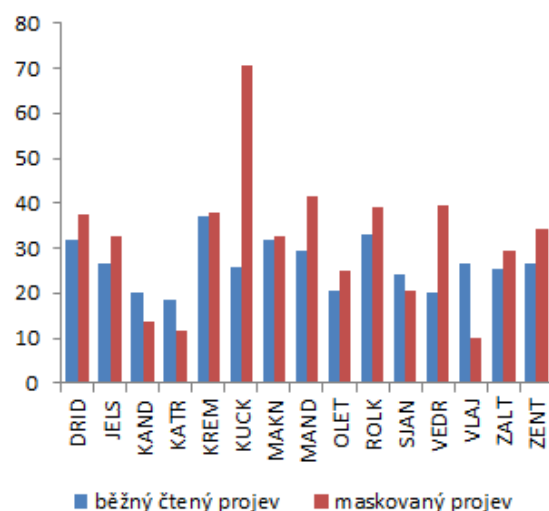
Tabulka 7-2: střední základní frekvence



Graf 7-5: střední  $f_0$

mluvčí	běžný projev	maskovaný projev	rozdíl
DRID	32	38	6
JELS	27	33	6
KAND	20	14	-6
KATR	18	12	-7
KREM	37	38	1
KUCK	26	71	45
MAKN	32	32	1
MAND	29	41	12
OLET	21	25	4
ROLK	33	39	6
SJAN	24	20	-4
VEDR	20	40	19
VLAJ	27	10	-17
ZALT	25	29	4
ZENT	27	34	7

Tabulka 7-3: směrodatné odchylky  $f_0$



Graf 7-6: porovnání velikosti směrodatných odchylek středních  $f_0$

### 7.2.2. Vokální formanty

Byla provedena analýza formantů F1, F2 a F3. Pomocí programu Praat byly extrahovány hodnoty formantů F1, F2 a F3 v ekvidistantních bodech v 10 ms intervalech pro krátké vokály /ɪ e a o u/ a dlouhé vokály /i: a: u:/. Nebyly analyzovány diftongy a také vokály /e:/ a /o:/, které se v textu, jež mluvčí předčítali maskovaným hlasem, neobjevily.

Následně byl proveden t-test pro hodnoty formantů extrahované z maskovaných a nemaskovaných projevů mluvčích. Na hladině významnosti  $\alpha = 0,05$  byla určena statistická významnost změn formantů, které se objevily u mluvčích mezi jejich maskovanými a nemaskovanými projevy.

Výsledky výpočtů představují tabulky 7-F1-1 až 7-F3-8 v příloze číslo 2. Červeně jsou vyznačeny p-hodnoty získané pomocí statistického testu nižší než stanovená hladina významnosti, na jejichž základě byl rozdíl mezi hodnotami daného formantu mluvčího při maskovaném a nemaskovaném projevu stanoven jako statisticky signifikantní.

Ukázalo se, že většina mluvčích je schopna prostřednictvím maskování hlasu statisticky významně modifikovat své vokální formanty, resp. formanty F1, F2 a F3.

Nejúspěšnější byli mluvčí při modifikaci formantu F1, jehož změna byla vyhodnocena jako statisticky významná v 82,5 % případů. (Tj. z posunů prvního formantu u všech vokálů v rámci projevů všech mluvčích bylo 82,5 % vyhodnoceno jako statisticky signifikantních. Např. změny F1 u vokálu /ɪ/ byly jakožto statisticky signifikantní vyhodnoceny u 10 mluvčích z 15, tj. u 67 %; ve zmíněných 82,5 % se pak jako statisticky významný ukázal posun F1 u všech analyzovaných vokálů, tzn. u 8 vokálů vyslovovaných 15 mluvčími.) Změny F2 byly vyhodnoceny jako statisticky signifikantní v 73,3 %, změny F3 pak v 67,5 % případů.

F1 byl v 67,5 % případů zvýšen, což svědčí o většinové otevřenější výslovnosti při maskování hlasu. V 68,3 % došlo ke zvýšení F2, mluvčí tedy měli častěji tendenci k posunům artikulace do anteriorní oblasti než naopak k faryngalizaci.

Statistický test však v několika případech vyhodnotil jako statisticky signifikantní také rozdíly mezi formanty v rámci jednoho typu projevu mluvčího, v tomto případě tedy tato metoda testování rozdílů mezi maskovaným a nemaskovaným projevem nebyla shledána efektivní.

### 7.2.3. Tempo

Bylo vyčísleno mluvní (MT) a artikulační tempo (AT) pro maskovaný i nemaskovaný projev každého mluvčího, vypočítán byl průměrný počet vyslovených slabik za sekundu ( $\sigma/s$ ).

Průměrné mluvní tempo (tj. vyčíslené pro celý řečový materiál, včetně pauz a dysfluencí) při běžném čteném projevu dosahovalo 4,57  $\sigma/s$ , v případě maskovaného projevu bylo v průměru sníženo na 4,10  $\sigma/s$ . O více než 1 slabiku za sekundu bylo oproti průměru vyšší mluvní tempo mluvčího ZENT, které dosáhlo 5,74  $\sigma/s$  při běžném a 5,51  $\sigma/s$  při maskovaném projevu. Naopak tempo pomalejší než 4  $\sigma/s$  vykazali při běžném projevu mluvčí MAND (3,69  $\sigma/s$ ) a MAKN (3,92  $\sigma/s$ ), přičemž mluvčí MAND hovořil pomaleji oproti průměrnému mluvnímu tempu i při projevu maskovaném (3,22  $\sigma/s$ ), zatímco rozdíl mluvního tempa mluvčího MAND a průměrného mluvního tempa při maskovaném projevu činil pouze 0,24  $\sigma/s$ . Naopak výrazněji nižší mluvní tempo oproti průměru (více než o 0,5  $\sigma/s$ ) bylo vyčísleno u mluvčích KAND (3,21  $\sigma/s$ ) a DRID (3,25  $\sigma/s$ ).

Hodnoty mluvního tempa mluvčích při maskovaném i nemaskovaném projevu ukazuje tabulka 7-4.

Průměrné artikulační tempo (tj. s vyloučením pauz a dysfluenčních prvků) při nemaskovaném projevu činilo 6,15  $\sigma/s$ , při projevu maskovaném pak bylo sníženo na 5,89  $\sigma/s$ . Výrazně vyšší artikulační tempo (o více než 1  $\sigma/s$ ) vykazal mluvčí KREM, a to při běžném (7,19  $\sigma/s$ ) i maskovaném projevu (7,35  $\sigma/s$ ). Vyšší artikulační tempo při

mluvčí	běžný projev	maskovaný projev	rozdíl
DRID	4,01	3,25	0,76
JELS	4,00	4,42	-0,43
KAND	4,73	3,21	1,52
KATR	4,07	3,94	0,13
KREM	5,10	4,48	0,62
KUCK	4,81	4,54	0,27
MAKN	3,92	3,86	0,06
MAND	3,69	3,22	0,48
OLET	4,19	4,35	-0,16
ROLK	4,94	4,22	0,71
SJAN	5,04	4,83	0,21
VEDR	4,48	3,67	0,80
VLAJ	5,29	3,95	1,34
ZALT	4,50	4,06	0,44
ZENT	5,74	5,51	0,23

**Tabulka 7-4: mluvní tempo**

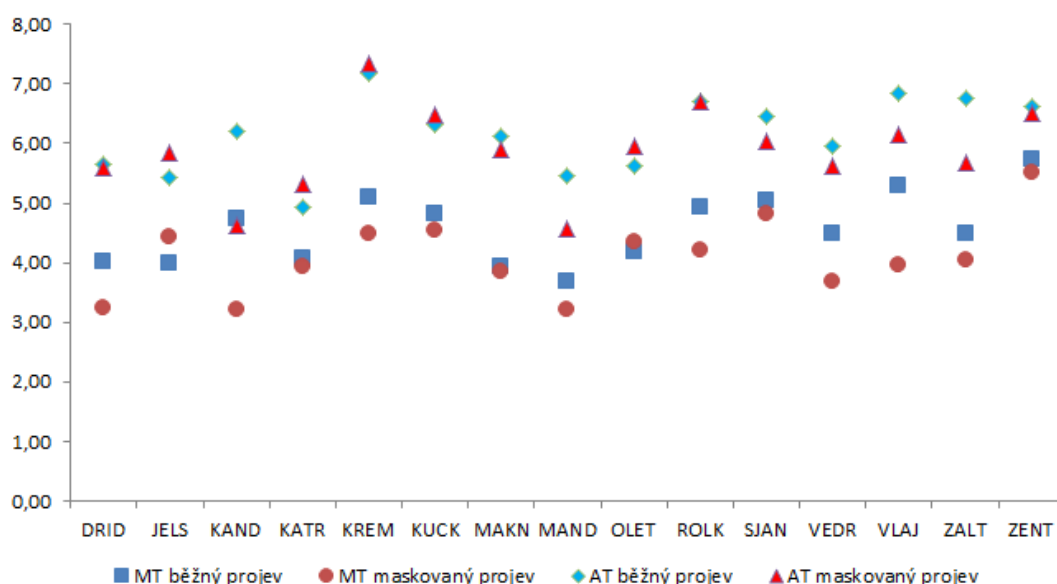
mluvčí	běžný projev	maskovaný projev	rozdíl
DRID	5,64	5,59	0,05
JELS	5,43	5,85	-0,42
KAND	6,21	4,62	1,59
KATR	4,93	5,32	-0,39
KREM	7,19	7,35	-0,16
KUCK	6,32	6,49	-0,18
MAKN	6,13	5,91	0,21
MAND	5,46	4,57	0,89
OLET	5,62	5,95	-0,33
ROLK	6,71	6,70	0,00
SJAN	6,45	6,03	0,42
VEDR	5,97	5,62	0,34
VLAJ	6,84	6,15	0,69
ZALT	6,77	5,67	1,09
ZENT	6,61	6,52	0,09

**Tabulka 7-5: artikulační tempo**



maskovaném projevu bylo zjištěno také u mluvčího ROLK (6,70  $\sigma/s$ ). Naopak nižší artikulační tempo při maskovaném projevu oproti průměru (o více než 1  $\sigma/s$ ) bylo nalezeno u mluvčích MAND (4,57) a KAND (4,62). Mluvčí MAND také, jak již bylo zmíněno, vykázal nižší mluvní tempo oproti průměru při maskovaném i nemaskovaném projevu, jeho artikulační tempo při nemaskovaném projevu však bylo naopak ve srovnání s průměrnou hodnotou vyšší (5,46  $\sigma/s$ , tj. vyšší než průměr o 0,69  $\sigma/s$ ). U mluvčího KAND bylo zjištěno nižší mluvní i artikulační tempo při maskovaném projevu, ne však při projevu běžném, při němž bylo naopak ve srovnání s průměrnou hodnotou mírně vyšší (o 0,16  $\sigma/s$  v případě mluvního a o 0,06 v případě artikulačního tempa).

Artikulační tempo mluvčích při běžném i maskovaném projevu vyčísluje tabulka 7-5, porovnání hodnot pak poskytuje graf 7-7.



**Graf 7-7:** porovnání mluvního a artikulačního tempa při maskovaném a nemaskovaném projevu

### 7.3. Percepční testy

#### 7.3.1. Percepční test č. 1

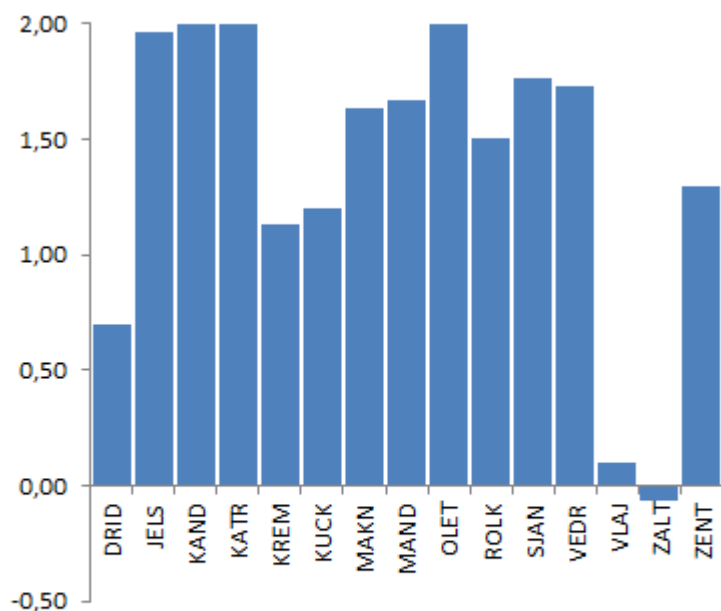
Pomocí prvního percepčního testu bylo zjišťováno, jak přirozeně zní maskovaný projev mluvčího, což bylo následně porovnáno s jeho běžným (nemaskovaným) projevem. Schopnost mluvčích maskovat svůj hlas tak, aby působil dojmem, že jde o přirozený projev, tj. aby posluchač nebyl schopen rozeznat, že se jedná o maskovanou

řeč, může být důležitým ukazatelem ve forenzní praxi – aby mohla být sporná nahrávka analyzována jakožto nahrávka maskovaného projevu mluvčího, musí být nejprve jako taková rozeznána. Pokud je mluvčí schopen efektivně maskovat rysy svého přirozeného projevu a zároveň hovořit tak, aby jeho hlas zněl přirozeně, může být značně zkomplikováno profilování mluvčího i proces jeho identifikace.

Z 15 mluvčích vybraných k akustické analýze a k percepčním testům bylo při poslechové analýze u 11 hodnocen maskovaný projev jako relativně přirozený.

Při poslechovém testu byli posluchači dotázáni, zda mají dojem, že mluvčí, který je zachycen na předložené nahrávce, maskuje svůj hlas. Odpovědi volili na pětibodové Likertově škále: „určitě ne“ (-2), „spíše ne“ (-1), „těžko říci“ (0), „spíše ano“ (+1), „určitě ano“ (+2). Na základě odpovědí respondentů bylo následně vypočítáno průměrné hodnocení maskovaného projevu každého mluvčího.

Respondenti ohodnotili maskovaný projev většiny mluvčích jako nepřirozený, výsledku mezi 0 a 2 dosáhlo 14 mluvčích, pouze u mluvčího ZALT se průměrné hodnocení dostalo do záporných hodnot (-0,07). Mezi 0 a 1 se nacházelo průměrné hodnocení mluvčích VLAJ (0,10) a DRID (0,7). Porovnání průměrných hodnocení projevů mluvčích představuje graf 7-8.

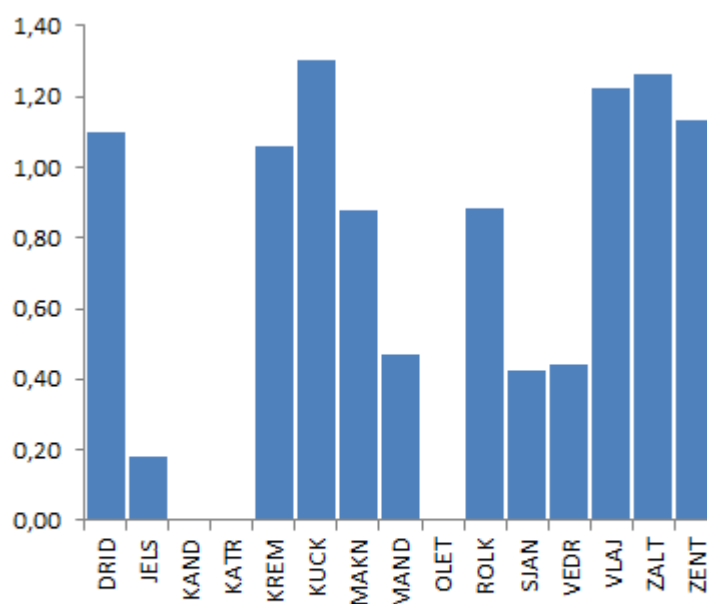


**Graf 7-8: průměrná hodnocení maskovaných projevů mluvčích**

Hodnocení posluchačů se u některých mluvčích shodovala, jednalo se o mluvčí, jejichž projev všichni respondenti ohodnotili jako zjevně maskovaný (+2): KAND, KATR a OLET, u těchto mluvčích se tedy směrodatná odchylka hodnocení rovnala 0 Hz. V hodnocení maskovaných projevů ostatních mluvčích se posluchači neshodovali, největší směrodatná odchylka hodnocení byla zaznamenána u mluvčích VLAJ (1,22 Hz), ZALT (1,26 Hz) a KUCK (1,30 Hz). Směrodatné odchylky hodnocení maskovaných projevů jednotlivých mluvčích ukazuje graf 7-9. Jak směrodatné odchylky, tak průměrná hodnocení pak vyčísluje tabulka 7-6.

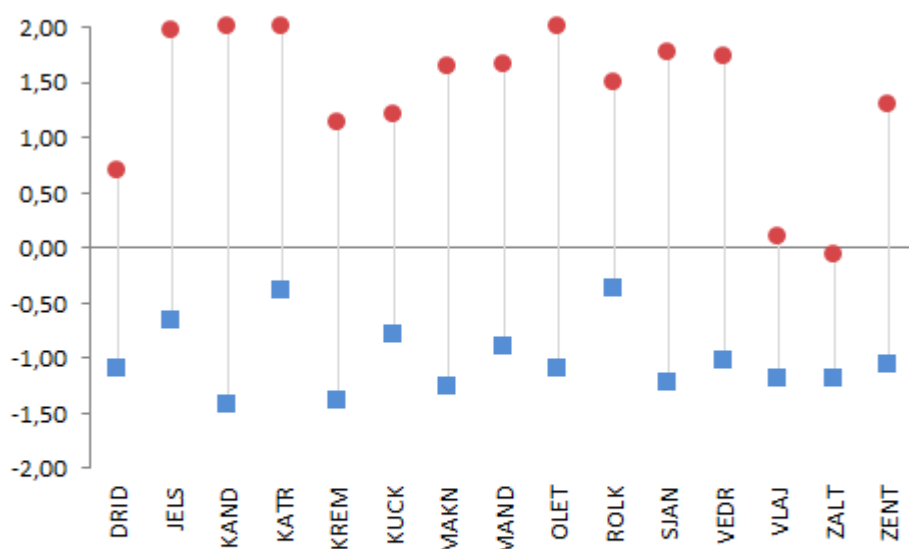
mluvčí	průměrné hodnocení	směrodatná odchylka
DRID	0,70	1,10
JELS	1,97	0,18
KAND	2,00	0,00
KATR	2,00	0,00
KREM	1,13	1,06
KUCK	1,20	1,30
MAKN	1,63	0,87
MAND	1,67	0,47
OLET	2,00	0,00
ROLK	1,50	0,89
SJAN	1,77	0,42
VEDR	1,73	0,44
VLAJ	0,10	1,22
ZALT	-0,07	1,26
ZENT	1,30	1,13

**Tabulka 7-6:** hodnocení maskovaných projevů mluvčích



**Graf 7-9:** směrodatné odchylky hodnocení maskovaných projevů mluvčích

Hodnocení přirozenosti bylo při tomto percepčním testu vynášeno také pro nemaskované projevy mluvčích, které sloužily jako srovnávací položky. Žádný z nemaskovaných projevů nebyl hodnocen jako určitě nemaskovaný (-2) všemi posluchači; nejvíce přirozeně zněl respondentům nemaskovaný hlas mluvčího KAND, jehož průměrné hodnocení bylo vyčísleno na -1,43; nejméně byli posluchači naopak přesvědčeni o přirozenosti hlasu mluvčího ROLK, průměrné hodnocení jeho nemaskovaného projevu bylo -0,37. Porovnání průměrných hodnocení maskovaných i nemaskovaných projevů předkládá graf 7-10, červené značky představují hodnoty získané pro maskované projevy mluvčích, modré pro projevy nemaskované.



**Graf 7-10: průměrná hodnocení maskovaných i nemaskovaných projevů mluvčích**

### 7.3.2. Percepční test č. 2

Druhý percepční test představoval modelovou rekognici. Respondentům byly postupně předkládány ukázky maskovaných projevů mluvčích (sporné nahrávky), k ukázce maskovaného projevu daného mluvčího byla vždy přiřazena nahrávka jeho nemaskovaného hlasu a nahrávky hlasů dvou figurantů (srovnávací nahrávky). Posluchači ze tří nemaskovaných položek volili tu, o níž se domnívali, že je na ní

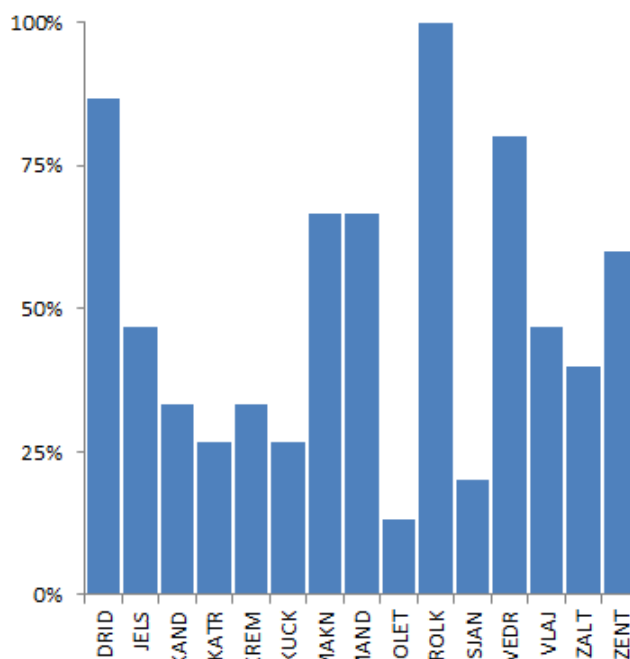
zachycen hlas téhož mluvčího jako na položce maskované. Pomocí tohoto percepčního testu byla zjišťována míra obtížnosti identifikace mluvčího.

Všichni respondenti na základě maskovaného projevu správně určili mluvčího ROLK. Nadpoloviční správnost identifikace byla zaznamenána také u mluvčích DRID (87 %), MAKN (67 %), MAND (67 %), VEDR (80 %) a ZENT (60 %). Ostatní mluvčí byli více než v polovině případů při rekognici zaměněni s jedním z figurantů.

Většina mluvčích z vybraného vzorku (9 z 15) dosáhla natolik efektivního maskování hlasu, že úspěšnost jejich rekognice byla méně než poloviční. Nejvyšší chybovost vykázali respondenti při identifikaci mluvčího OLET, jež se podařilo správně určit pouze dvěma z nich (tj. 13 %). Úspěšnost rekognice pod 50 % byla dále pozorována u mluvčích JELS (47 %), KAND (33 %), KATR (27 %), KREM (33 %), KUCK (27 %), SJAN (20 %), VLAJ (47 %) a ZALT (40 %).

Míru úspěšnosti rekognice pro jednotlivé mluvčí představuje graf 7-11.

Co se týče chybných identifikací, nejčastěji byl spojen nemaskovaný projev mluvčího KAND s maskovaným projevem mluvčího KATR (67 %) a nemaskovaný projev mluvčího ROLK s maskovaným projevem mluvčího KREM (67 %). Více než v polovině případů také posluchači spojili maskovaný projev mluvčího OLET s nemaskovaným



**Graf 7-11:** úspěšnost rekognice u jednotlivých mluvčích

mluvčí	figurant 1 (případů záměny)	figurant 2 (případů záměny)
DRID	MAKN (7 %)	OLET (7 %)
JELS	KREM (27 %)	VLAJ (27 %)
KAND	KATR (27 %)	SJAN (40 %)
KATR	KAND (67 %)	MAND (7 %)
KREM	DRID (0 %)	ROLK (67 %)
KUCK	VEDR (27 %)	ZALT (47 %)
MAKN	DRID (33 %)	OLET (0 %)
MAND	VEDR (13 %)	ZENT (20 %)
OLET	KAND (60 %)	KUCK (27 %)
ROLK	KUCK (0 %)	JELS (0 %)
SJAN	MAKN (60 %)	ZALT (20 %)
VEDR	VLAJ (20 %)	ZENT (0 %)
VLAJ	JELS (33 %)	MAND (20 %)
ZALT	KATR (0 %)	ROLK (47 %)
ZENT	KREM (13 %)	SJAN (27 %)

**Tabulka 7-7:** figuranti přiřazení k jednotlivým mluvčím a míra chybné identifikace

projevem mluvčího KAND (60 %) a maskovaný projev mluvčího SJAN s nemaskovaným projevem mluvčího MAKN (60 %).

Figuranty přiřazené k jednotlivým mluvčím při rekognici a míru chybné identifikace pro každého z nich zachycuje tabulka 7-7.

## **8. Souhrn výsledků pro jednotlivé mluvčí**

### **8.1. Mluvčí DRID**

Při poslechové analýze byl dojem, kterým působil maskovaný projev tohoto mluvčího, hodnocen jako přirozený, obtížnost jeho rozpoznání byla vyhodnocena jako střední. Bylo zaznamenáno zvýšení základní frekvence a užití dyšné fonace. Nebyl pozorován rozdíl v tempu řeči.

Akustická analýza odhalila nárůst střední  $f_0$  o 31 Hz (5,6 půltónu) při maskovaném projevu oproti projevu čtenému běžným způsobem. Variabilita  $f_0$  v rámci maskovaného a nemaskovaného projevu se výrazně nelišila: směrodatná odchylka  $f_0$  při běžném projevu dosáhla 32 Hz, při projevu maskovaném pak 38 Hz.

Statisticky signifikantní rozdíly formantů při běžném a maskovaném projevu byly u tohoto mluvčího zjištěny v 63 % případů. F1 byl u 87,5 % zvýšen a u 62,5 % vokálu snížen, což nasvědčuje otevřenější a zadnější artikulaci.

Mluvní tempo mluvčího DRID při běžném projevu dosáhlo 4,01  $\sigma/s$ , zatímco při projevu maskovaném kleslo na 3,25  $\sigma/s$ . Byl zaznamenán také pokles artikulačního tempa, avšak pouze o 0,05  $\sigma/s$  (z 5,64  $\sigma/s$  při běžném na 5,59  $\sigma/s$  při maskovaném projevu).

Při percepčním testu č. 1, kdy respondenti hodnotili, zda jim připadá, že mluvčí maskuje svůj hlas na škále od -2 do 2 (-2 určitě ano, -2 určitě ne), jeho maskovaný projev v průměru dosáhl skóre 0,7, přičemž respondenti se ve svých hodnoceních lišili – směrodatná odchylka této hodnoty činila 1,1.

V percepčním testu č. 2, tj. při rekognici, byl mluvčí DRID správně identifikován v 86 % případů.

### **8.2. Mluvčí JELS**

Maskovaný projev mluvčího JELS byl při poslechové analýze vyhodnocen jako znějící nepřirozeně, jeho identifikovatelnost byla určena jako obtížná. Bylo pozorováno zvýšení základní frekvence, využití nazalizace a změny artikulace. Mluvčí využil modální fonaci a nebyla zaregistrována změna tempa.

Pomocí akustické analýzy byl vyčíslen rozdíl mezi střední  $f_0$  při maskovaném a nemaskovaném projevu, který činil 59 Hz (střední  $f_0$  byla rovna 169 Hz při nemaskovaném a 228 Hz při maskovaném projevu, jednalo se o nárůst o 5,2 půltónu).

Směrodatná odchylka při nemaskovaném projevu činila 27 Hz a při maskovaném 33 Hz, variabilita  $f_0$  tedy při maskování nebyla výrazně změněna.

Své vokalické formanty tento mluvčí při maskování hlasu statisticky významně změnil v 92 % případů. U 87,5 % vokálů došlo ke zvýšení F1 a u 62,5 % vokálů ke zvýšení F2, artikulace vokálů tedy byla častěji otevřenější a přednější.

Mluvní tempo tohoto mluvčího bylo při běžném projevu vyčísleno na 4  $\sigma/s$ , zatímco při projevu maskovaném činilo 4,42  $\sigma/s$ . Tempo artikulační dosáhlo při běžném projevu 5,43  $\sigma/s$  a při maskovaném projevu 5,85  $\sigma/s$ . Mluvní i artikulační tempo tedy při maskování hlasu vzrostlo.

Respondenti při percepčním testu č. 1 ohodnotili maskovaný projev mluvčího JELS jako nepřírozený průměrným skóre 1,97, přičemž mezi respondenty panovala povětšinou shoda – směrodatná odchylka hodnocení byla pouze 0,18.

Maskování mluvčího JELS bylo vyhodnoceno jako nápadné, bylo však relativně efektivní – při rekognici byl tento mluvčí úspěšně identifikován pouze ve 46 % případů.

### **8.3. Mluvčí KAND**

Při poslechové analýze byl dojem, kterým působil maskovaný hlas mluvčího KAND, vyhodnocen jako nepřírozený, jeho identifikovatelnost na základě maskovaného projevu byla shledána obtížnou. Bylo zaregistrováno snížení  $f_0$ , využití třepené fonace, zmenšení variability  $f_0$  (umělá monotónnost) a změna artikulace.

Pomocí akustické analýzy byla vypočtena střední  $f_0$  při maskovaném a nemaskovaném projevu; při nemaskovaném dosáhla 115 Hz, při maskovaném klesla o 29 Hz (5 půltónů) na 86 Hz. Navzdory výsledkům zaznamenaným při poslechové analýze nebylo zjištěno výrazné snížení variability  $f_0$ , směrodatná odchylka klesla pouze o 6 Hz z 20 Hz na 14 Hz.

Změny jeho formantů byly vyhodnoceny jako statisticky významné v 75 % případů. U všech vokálů došlo ke zvýšení F1, mluvčí při maskování hlasu tedy zvolil otevřenější artikulaci. Poměr zvýšení a snížení F2 byl totožný.

Mluvní tempo mluvčího KAND bylo při běžném projevu rovno 4,73  $\sigma/s$ , při projevu maskovaném vzrostlo na 3,21  $\sigma/s$ . Artikulační tempo však naopak výrazně kleslo: z 6,21  $\sigma/s$  na 4,62  $\sigma/s$ . Mluvčí tedy při maskování hlasu pomaleji artikuloval, avšak produkoval méně pauz a dysfluencí.



V rámci percepčního testu č. 1 byl jeho maskovaný projev všemi posluchači hodnocen jako nepřirozený, všichni respondenti na otázku, zda mluvčí maskuje svůj hlas, zvolili odpověď „určitě ano“, jeho skóre tedy dosáhlo 2 se směrodatnou odchylkou 0.

Stejně jako v případě mluvčího JELS se i způsob maskování, jež zvolil mluvčí KAND, navzdory nepřirozenému dojmu ukázal jako efektivní; při rekognici v rámci percepčního testu č. 2 byl úspěšně identifikován pouze 33 % respondentů.

#### **8.4. Mluvčí KATR**

U tohoto mluvčího bylo v rámci poslechové analýzy při maskování hlasu pozorováno snížení  $f_0$  a změna artikulace, mluvčí využil modální fonaci a nebyla zaregistrována změna tempa. Jeho identifikovatelnost byla hodnocena jako obtížná a dojem, jímž působil jeho maskovaný projev, byl určen jako přirozený.

Střední  $f_0$  vyčíslená pomocí akustické analýzy činila 113 Hz při běžném a 112 Hz při maskovaném projevu, nedošlo tedy k její změně. Směrodatná odchylka  $f_0$  dosáhla 18 Hz při běžném a 12 Hz při maskovaném projevu, došlo tedy k určitému snížení variability  $f_0$ .

Statisticky signifikantní změnu formantů provedl tento mluvčí v 79 % případů. U všech vokálů došlo ke snížení  $F_1$ , mluvčí se tedy uchýlil k zavřenější artikulaci. Zvýšení i snížení  $F_2$  dosáhlo 50 %.

Při nemaskovaném projevu hovořil mluvčí KATR mluvním tempem 4,07  $\sigma/s$ , v případě maskovaného projevu kleslo na 3,94. Artikulační tempo naopak vzrostlo, a to z 4,93  $\sigma/s$  při běžném projevu na 5,32  $\sigma/s$  při projevu maskovaném. Při maskovaném projevu tedy mluvčí artikuloval rychleji, objevovalo se v něm však více pauz či dysfluencí.

Stejně jako mluvčí KAND byl také mluvčí KATR při percepčním testu č. 1 odhalen jako maskující svůj hlas všemi respondenty.

Úspěšnost jeho identifikace byla ještě nižší, než tomu bylo v případě mluvčího KAND – dosáhla pouze 26 %. Zvolený způsob maskování hlasu se tedy ukázal jako efektivní.

#### **8.5. Mluvčí KREM**

Při poslechové analýze byl jeho maskovaný projev shledán přirozeným a jeho identifikace středně obtížnou. Bylo pozorováno zvýšení  $f_0$  a užití dyšné fonace.

Navzdory výsledkům poslechové analýzy nebyl pomocí akustické analýzy odhalen rozdíl ve střední  $f_0$  při maskování hlasu – při přirozeném projevu byla střední  $f_0$

152 Hz, při projevu maskovaném 153 Hz. Ani variabilita  $f_0$  se neukázala být výrazně modifikovanou, směrodatná odchylka činila při běžném projevu 37 Hz, při maskovaném pak 38 Hz.

Změny formantů při maskování hlasu byly u tohoto mluvčího vyhodnoceny jako statisticky signifikantní v 58 % případů. Zvýšení i snížení  $F_1$  bylo stejně časté, u 75 % vokálů došlo ke zvýšení  $F_2$ , mluvčí tedy při maskování hlasu zvolil přednější artikulaci vokálů.

Mluvní tempo tohoto mluvčího činilo při běžném čteném projevu 5,1  $\sigma/s$ , při projevu maskovaném bylo sníženo na 4,48  $\sigma/s$ . Artikulační tempo naopak vzrostlo ze 7,19 na 7,35  $\sigma/s$ .

Při percepčním testu č. 1 byl maskovaný hlas mluvčího KREM hodnocen jako spíše nepřirozený, skóre dosáhlo 1,13. Respondenti se ve svých hodnoceních poměrně značně rozcházeli, směrodatná odchylka činila 1,06.

Při rekognici byl mluvčí KREM úspěšně identifikován ve 43 % případů.

## **8.6. Mluvčí KUCK**

Mluvčí KUCK při maskování hlasu zvolil šepot. Změna tempa nebyla zaznamenána. Dojem maskovaného projevu byl vyhodnocen jako nepřirozený, rozpoznatelnost byla shledána obtížnou.

Při akustické analýze byla střední  $f_0$  při přirozeném projevu vyčíslena na 157 Hz, při projevu maskovaném na 174 Hz (nárůst o 1,8 půltónu), výsledky analýzy ukázaly nárůst směrodatné odchylky z 26 Hz na 71 Hz. Výpočet nebyl manuálně korigován, v tomto případě zřejmě výsledek akustické analýzy není příliš vypovídající, vzhledem ke skutečnosti, že mluvčí KUCK jako strategii pro maskování hlasu zvolil šepot.

Rozdíly mezi vokalickými formanty při maskovaném a nemaskovaném projevu byly u tohoto mluvčího vyhodnoceny jako statisticky významné v 75 % případů. U všech vokálů bylo zaznamenáno zvýšení  $F_1$ , u 75 % vokálů pak také zvýšení  $F_2$ , což vypovídá o otevřenější a přednější artikulaci vokálů u tohoto mluvčího.

Mluvní tempo činilo 4,81  $\sigma/s$  při běžném a 4,54  $\sigma/s$  při maskovaném projevu, kleslo tedy o 0,13  $\sigma/s$ . Artikulační tempo bylo při běžném projevu 6,32  $\sigma/s$ , při maskovaném pak 6,49  $\sigma/s$ , na rozdíl od mluvního tempa tedy došlo k jeho nárůstu, a to o 0,18  $\sigma/s$ .

V rámci percepčního testu č. 1 byl jeho maskovaný projev hodnocen jako spíše nepřirozený s poměrně vysokou směrodatnou odchylkou (1,3), zřejmě zejména kvůli

skutečnosti, že respondenti si nebyli jisti, zda šepot vyhodnotit jako způsob maskování hlasu.

Při percepčním testu č. 2 byla úspěšnost identifikace mluvčího KUCK 26%.

### **8.7. Mluvčí MAKN**

Stejně jako mluvčí KUCK zvolil i mluvčí MAKN při maskování šepot. Ani u něj nebyla pozorována změna tempa. Jeho maskovaný projev byl vyhodnocen jako nepřirozený a rozpoznatelnost jako obtížnější.

Při akustické analýze byla střední  $f_0$  při nemaskovaném projevu vyčíslena na 139 Hz se směrodatnou odchylkou 32 Hz. Při maskovaném projevu byla stanovena na 301 Hz s touž směrodatnou odchylkou, ovšem stejně jako v případě předchozího mluvčího není výsledek extrakce  $f_0$  vypovídající vzhledem k šepotu, jehož mluvčí při maskování hlasu využil.

Statisticky signifikantní změny byly u tohoto mluvčího nalezeny v 96 % případů. U všech vokálů došlo ke zvýšení  $F_1$  i  $F_2$ , tento mluvčí tedy při maskování zvolil otevřenější a přednější artikulaci.

Mluvní tempo při běžném projevu dosáhlo 3,92  $\sigma/s$ , při maskovaném projevu pak 3,86  $\sigma/s$ . Artikulační tempo bylo při běžném projevu 6,13  $\sigma/s$ , při maskovaném projevu kleslo na 5,91  $\sigma/s$ .

V rámci percepčního testu č. 1 byl maskovaný projev mluvčího MAKN vyhodnocen jako spíše nepřirozený s výsledným průměrným skóre 1,63 a směrodatnou odchylkou 0,87. Jeho projev byl hodnocen jako nápadněji maskovaný než projev mluvčího KUCK, ačkoli oba mluvčí jakožto strategii při maskování hlasu zvolili šepot.

Na rozdíl od mluvčího KUCK byl při druhém percepčním testu mluvčí MAKN úspěšně rozpoznán ve většině případů, správně jej na základě hlasu identifikovalo 67 % respondentů.

### **8.8. Mluvčí MAND**

Při poslechové analýze bylo u maskovaného projevu tohoto mluvčího zaznamenáno zvýšení základní frekvence a imitace sigmatismu. Fonace byla zvolena modální a tempo nedoznalo výrazných změn. Dojem jeho maskovaného projevu byl vyhodnocen jako přirozený a rozpoznatelnost jako obtížná.

Při akustické analýze byl odhalen nárůst střední  $f_0$  o 28 Hz (3,3 půltónu) ze 132 Hz při nemaskovaném projevu na 160 Hz při projevu maskovaném. Bylo zaznamenáno zvětšení variability  $f_0$ , směrodatná odchylka vzrostla z 29 na 41.

Posuny vokálních formantů mezi maskovaným a nemaskovaným projevem byly vyhodnoceny jako statisticky signifikantní v 58 % případů. U 87,5 % vokálů bylo zaznamenáno zvýšení  $F_1$ , u 75 % vokálů pak také zvýšení  $F_2$ , což vypovídá o otevřenější a přednější artikulaci při maskování hlasu.

Mluvní tempo při nemaskovaném projevu činilo 3,69  $\sigma/s$ , při maskovaném projevu kleslo na 3,22  $\sigma/s$ . Artikulační tempo pak kleslo z 5,46  $\sigma/s$  na 4,57  $\sigma/s$ .

Při percepčním testu č. 1 byl jeho maskovaný projev respondenty ve většině případů hodnocen jako nepřirozený, průměrné hodnocení dosáhlo 1,67 se směrodatnou odchylkou 0,47.

V rámci rekognice byl mluvčí MAND správně identifikován v 67 % případů.

## **8.9. Mluvčí OLET**

Maskovaný projev mluvčího OLET byl při percepční analýze vyhodnocen jako přirozený a jeho rozpoznání bylo shledáno obtížným. Bylo zaznamenáno zvýšení základní frekvence a modifikace artikulace. Fonace byla zvolena modální a tempo nebylo výrazně změněno.

Akustická analýza ukázala zvýšení střední  $f_0$  o 35 Hz (4,7 půltónu) z původních 111 Hz na 146 Hz při maskovaném projevu. Variabilita  $f_0$  nebyla výrazně modifikována, směrodatná odchylka vzrostla z 21 Hz na 25 Hz.

Statisticky významné posuny vokálních formantů byly u tohoto mluvčího nalezeny v 83 % případů. U všech vokálů došlo k poklesu  $F_1$ , u 87,5 % pak i k poklesu  $F_2$ , tento mluvčí tedy při maskování svého hlasu zvolil zavřenější a zadnější artikulaci vokálů.

Při nemaskovaném projevu bylo mluvní tempo 4,19  $\sigma/s$ , při maskovaném projevu vzrostlo na 4,35. Artikulační tempo vzrostlo z 5,62  $\sigma/s$  na 5,95  $\sigma/s$ .

Výsledky percepčního testu č. 1 ukázaly, že všichni respondenti hodnotili maskovaný projev mluvčího OLET jako nepřirozený, jeho průměrné skóre dosáhlo 2 se směrodatnou odchylkou 0. Tyto výsledky byly v rozporu s poslechovou analýzou, při níž bylo shledáno, že daný řečový projev by mohl odpovídat běžnému hlasu mluvčího.

Strategie, kterou tento mluvčí zvolil při maskování hlasu, se ukázala být nejefektivnější ze všech analyzovaných, úspěšnost identifikace mluvčího OLET při

rekognici dosáhla pouhých 13 % (tj. úspěšně jej identifikovali pouze 2 respondenti z 15).

#### **8.10. Mluvčí ROLK**

Podle poslechové analýzy mluvčí ROLK při maskování hlasu zvolil dyšnou fonaci a imitaci cizineckého přízvuku, pravděpodobně ruského. Nebyla pozorována změna  $f_0$  ani tempa. Rozpoznání bylo vyhodnoceno jako obtížnější a maskovaný projev působil přirozeným dojmem.

Akustická analýza ukázala snížení střední  $f_0$  ze 140 Hz při běžném projevu na 124 Hz při maskovaném projevu (tj. pokles o 2,1 půltónu). Variabilita  $f_0$  mírně vzrostla, směrodatná odchylka byla zvýšena z 33 Hz na 39 Hz.

Rozdíly mezi vokálními formanty v rámci maskovaného a nemaskovaného projevu byly vyhodnoceny jako statisticky signifikantní v 58 % případů. U 75 % vokálů došlo k nárůstu  $F_1$ , u 87,5 % také k nárůstu  $F_2$ . Byla tedy využita otevřenější a přednější artikulace.

Mluvní tempo kleslo ze 4,94  $\sigma/s$  na 4,22  $\sigma/s$ , artikulační tempo zůstalo prakticky nezměněno; při nemaskovaném i maskovaném projevu činilo 6,7  $\sigma/s$ .

Maskovaný projev mluvčího ROLK byl respondenty vyhodnocen jako spíše nepřirozený, průměrné hodnocení dosáhlo 1,5. Posluchači se v hodnocení lišili, směrodatná odchylka činila 0,89.

Navzdory výsledkům poslechové analýzy se způsob maskování zvolený tímto mluvčím neukázal být efektivním, při rekognici jeho maskovaný projev s nemaskovaným správně spojilo všech 15 respondentů.

#### **8.11. Mluvčí SJAN**

Při poslechové analýze bylo u maskovaného projevu mluvčího SJAN pozorováno zvýšení základní frekvence a změna artikulace (posun jazyka dopředu při artikulaci vokálů). Fonace byla využita modální, tempo nebylo výrazně odlišné. Obtížnost jeho identifikace na základě maskovaného projevu byla vyhodnocena jako střední a jeho maskovaný hlas zněl přirozeně.

Podle akustické analýzy vzrostla při maskování hlasu střední základní frekvence ze 107 Hz na 156 Hz (došlo tedy ke zvýšení  $f_0$  o 6,5 půltónu). Variabilita  $f_0$  byla mírně redukována, směrodatná odchylka byla snížena z 24 Hz na 20 Hz.

Statisticky signifikantní změny formantů byly nalezeny v 75 % případů. U všech vokálů bylo zaznamenáno zvýšení F1 i F2, což vypovídá o otevřenější a přednější artikulaci.

Mluvní tempo tohoto mluvčího při běžném čteném projevu činilo 5,04  $\sigma/s$ , při maskovaném projevu bylo sníženo na 4,83  $\sigma/s$ . Artikulační tempo bylo sníženo z 6,45  $\sigma/s$  na 6,15  $\sigma/s$ .

Při prvním percepčním testu byl projev mluvčího SJAN hodnocen jako nepřirozený, průměrné hodnocení dosáhlo 1,77 se směrodatnou odchylkou 0,42.

Při druhém percepčním testu byl mluvčí SJAN úspěšně identifikován ve 20 % případů.

### **8.12. Mluvčí VEDR**

U maskovaného projevu mluvčího VEDR byl při poslechové analýze zaznamenán nárůst základní frekvence a využití dyšné fonace. Změny tempa nebyly zaregistrovány. Jeho maskovaný hlas působil přirozeným dojmem a obtížnost jeho identifikace byla vyhodnocena jako střední.

Podle výsledků akustické analýzy vzrostla střední  $f_0$  u tohoto mluvčího ze 114 Hz na 169 Hz (tj. o 6,8 půltónu). Variabilita  $f_0$  byla zvýšena, její směrodatná odchylka byla v případě běžného projevu rovna 20 Hz, při maskovaném projevu pak 40 Hz.

Posuny vokalických formantů byly vyhodnoceny jako statisticky signifikantní v 54 % případů. U 87,5 % vokálů vzrostl F1, u 62,5 % pak také F2. Tento mluvčí měl tedy tendenci se uchýlovat k otevřenější a přednější artikulaci vokálů.

Mluvní tempo při maskování hlasu kleslo o 0,81  $\sigma/s$  ze 4,48  $\sigma/s$  na 3,67  $\sigma/s$ , artikulační tempo pak pouze o 0,34  $\sigma/s$  z 5,97  $\sigma/s$  na 5,62  $\sigma/s$ .

Při prvním percepčním testu dosáhl maskovaný projev tohoto mluvčího skóre 1,73 se směrodatnou odchylkou 0,44, většinou respondentů byl tedy hodnocen jako nepřirozený, resp. zjevně maskovaný.

Při rekognici byl mluvčí VEDR správně identifikován 80 % mluvčích.

### **8.13. Mluvčí VLAJ**

Podle výsledků poslechové analýzy tento mluvčí při maskování hlasu zvolil dyšnou fonaci a snížil svou základní frekvenci. Dojem, jímž působil jeho maskovaný hlas, byl shledán přirozeným, a obtížnost identifikace mluvčího byla vyhodnocena jako střední.

Při svém běžném projevu hovořil hlasem o střední základní frekvenci 133 Hz, Při maskování hlasu došlo k jejímu snížení na 99 Hz (pokles o 5,1 pultónu). Byla redukována variabilita  $f_0$ , směrodatná odchylka byla snížena z 27 Hz na 10 Hz.

U tohoto mluvčího byly pozorovány statisticky významné změny vokálních formantů v 67 % případů. U všech vokálů došlo k poklesu F1, u 87,5 % vokálů pak k nárůstu F2, mluvčí tedy zvolil zavřenější a přednější artikulaci.

Mluvní tempo bylo při maskování hlasu sníženo z 5,29  $\sigma/s$  na 3,95  $\sigma/s$ , artikulační tempo při maskování hlasu pokleslo z 6,84  $\sigma/s$  na 6,15  $\sigma/s$ .

V prvním percepčním testu se jeho skóre blížilo 0, respondenti si tedy v mnoha případech nebyli jisti, zda se jedná o maskovaný projev, či nikoliv. Průměrné hodnocení jeho maskovaného projevu dosáhlo 0,1. Posluchači se ve svých odpovědích rozcházel, směrodatná odchylka tohoto hodnocení dosáhla 1,22.

Při druhém percepčním testu dosáhla úspěšnost identifikace mluvčího VLAJ 47 %.

#### **8.14. Mluvčí ZALT**

Maskovaný projev mluvčího ZALT byl při poslechové analýze shledán přirozeně působícím, obtížnost jeho identifikace byla vyhodnocena jako střední. Bylo zaznamenáno zvýšení  $f_0$  a využití dyšné fonace. Nebyla pozorována změna tempa.

Akustická analýza ukázala zvýšení střední  $f_0$  při maskování hlasu ze 116 Hz na 150 Hz (nárůst o 4,5 pultónu), variabilita  $f_0$  mírně vzrostla, směrodatná odchylka stoupla z 25 Hz na 29 Hz.

V 63 % případů docílil tento mluvčí statisticky významných změn svých vokálních formantů. U všech vokálů byl zvýšen F1, u 62,5 % pak také F2, mluvčí ZALT se tedy uchýloval k otevřenější a přednější artikulaci.

Mluvní tempo kleslo při maskovaném projevu ze 4,50  $\sigma/s$  na 4,06  $\sigma/s$ , výrazněji kleslo tempo artikulační, a to z 6,77  $\sigma/s$  na 5,67  $\sigma/s$ .

Maskovaný projev mluvčího ZALT byl posluchači v rámci prvního percepčního testu jako jediný vyhodnocen jako spíše přirozený, ačkoliv jen mírně – jeho průměrné skóre bylo -0,07. Mezi respondenty v hodnocení nepanovala shoda, směrodatná odchylka činila 1,26.

Zároveň se způsob maskování zvolený tímto mluvčím také ukázal jako relativně efektivní, úspěšnost jeho identifikace dosáhla 53 %.

### 8.15. Mluvčí ZENT

U tohoto mluvčího bylo v rámci poslechové analýzy zaznamenáno snížení základní frekvence a užití dyšné fonace. Působení jeho projevu bylo vyhodnoceno jako přirozené a jeho identifikovatelnost byla shledána obtížnou.

Při akustické analýze byl zaznamenán pokles střední  $f_0$  pouze o 14 Hz (2 půltóny), a to ze 129 Hz na 115 Hz. Byla zvýšena variabilita  $f_0$ , směrodatná odchylka vzrostla z 27 Hz na 34 Hz.

Statisticky signifikantní posuny formantů byly pozorovány v 58 % případů. U 62,5 % vokálů došlo k poklesu  $F_1$ , u 87,5 % vokálů k nárůstu  $F_2$ , což vypovídá o zavřenější a přednější artikulaci při maskování hlasu.

Mluvní tempo mluvčího ZENT bylo při maskování hlasu sníženo z 5,74  $\sigma/s$  na 5,51  $\sigma/s$ , artikulační tempo pak mírně kleslo – z 6,61  $\sigma/s$  na 6,52  $\sigma/s$ .

Při percepčním testu č. 1 byl jeho maskovaný projev posluchači shledán spíše nepřirozeným, průměrné hodnocení dosáhlo 1,3. Respondenti se ve svém hodnocení jeho maskovaného projevu rozcházel, směrodatná odchylka dosáhla 1,13.

Tento mluvčí byl při modelové rekognici v rámci druhého percepčního testu úspěšně rozpoznán 60 % respondentů.



## 9. Diskuse

V rámci této práce bylo podrobeno analýze 80 nahrávek mluvčích, kteří libovolnými způsoby maskovali svůj hlas.

Ukázalo se, že pokud jsou mluvčí instruováni, aby měnili svůj hlas libovolným způsobem, většina z nich se rozhodne pro změnu základní frekvence, v analyzovaném vzorku mluvčích se ke změně tohoto parametru uchýlilo 71,25 % mluvčích. Z toho 42,5% mluvčích svou základní frekvenci zvýšilo a 28,75 % ji naopak snížilo. Tato skutečnost se shoduje s Künzelovým (2000) poznatkem, že při maskování hlasu většina mluvčí zvyšuje svou  $f_0$ , následuje snižování  $f_0$  a šepot.

Výrazně méně mluvčích se rozhoduje pro modifikace jiných parametrů, jakými jsou fonace či tempo; u modální fonace při maskování hlasu zůstalo 67,5 % mluvčích a tempo bez percepčně zaznamenanatelného rozdílu si zachovalo 87,5 % mluvčích.

V rámci maskování hlasu se objevovala např. nazalizace, denazalizace pomocí ucpaného nosu, chrapot, imitace řečových vad či napodobování cizineckého přízvuku, jednalo se však pouze o strategie volené jednotlivci, ačkoliv podle závěrů, které uvádí Künzel (2000), jsou imitace přízvuku a denazalizace relativně často využívanými strategiemi při maskování hlasu.

Mnoho mluvčích provedlo pouze výraznou změnu jednoho parametru, k témuž závěru dospěl také Masthoff (1996), který tuto skutečnost dává do souvislosti s potřebou zároveň udržet maskovací strategii a předat informaci, přičemž mluvčí musí vynakládat poměrně značné kognitivní úsilí.

Několik mluvčích zvolilo komplexnější strategie a projevy těchto jedinců byly podrobeny detailnější analýze – byla provedena akustická analýza jejich maskovaných i nemaskovaných projevů a na základě pořízených nahrávek byly sestaveny percepční testy.

Akustická analýza potvrdila, že při maskování hlasu dochází u většiny mluvčích ke snížení mluvního i artikulačního tempa; ke snížení průměrného mluvního tempa došlo u 13 z 15 sledovaných mluvčích, ke snížení tempa artikulačního pak u 10 z nich. Snížení artikulačního tempa může napovídat vyvíjenému kognitivnímu úsilí, pokud mluvčí čte a zároveň usiluje o udržení zvoleného způsobu maskování hlasu. Pokles mluvního tempa je následkem většího množství produkovaných pauz či dysfluencí. Tento výsledek však může být částečně ovlivněn také obsahem textu, jež mluvčí četli maskovaným hlasem, jednalo se totiž o jednu stranu telefonního hovoru – u mluvčích se tedy mohly objevovat delší odmlky na místech, kde by se nacházely repliky druhého participanta dialogu.

Při analýze vokálních formantů bylo pomocí dvouvýběrového t-testu nezávislých výběrů zjištěno, že většina mluvčích vykazuje při maskování hlasu statisticky signifikantní změny vokálních formantů F1, F2 a F3. Využitý statistický test však jako statisticky významné hodnotil i rozdíly mezi vokálními formanty v rámci jednoho typu projevu daného mluvčího, využitá metoda se tedy neukázala jako vhodná pro tuto analýzu, resp. pro fonetickou identifikaci mluvčího. Tato skutečnost mohla být způsobena malým objemem zkoumaného materiálu, přesněji řečeno nízkým počtem vokálů, které se objevovaly ve čtených textech. Lepší informace by mohla poskytnout analýza většího množství dat. Při akustické analýze sporných nahrávek ve forenzní praxi však bývá často k dispozici omezené množství řečového materiálu, je tedy otázkou, zda je analýza vokálních formantů jednotlivých vokálů v takových případech vhodným nástrojem.

Pomocí percepčních testů bylo ověřeno, zda maskovaný projev daného mluvčího působí přirozeným dojmem, a do jaké míry je obtížná identifikace mluvčího na základě jeho maskovaného projevu.

Pokud maskovaný projev mluvčího působí přirozeným dojmem, tj. tak, že by se mohlo jednat o běžný projev mluvčího, se kterým posluchač není obeznámen, může ve forenzní praxi dojít ke komplikaci zužování okruhu podezřelých. Pokud vyšetřovatel nepozná, že je hlas na pořízené sporné nahrávce záměrně modifikovaný, bude pátráno po pachateli s jiným hlasovým projevem, než jaký ve skutečnosti má. Toto může také zkomplikovat profilování mluvčího – např. pokud pachatel úspěšně imituje cizinecký přízvuk.

V analyzovaném řečovém vzorku se však objevilo jen málo mluvčích, jejichž maskovaný projev působil přirozeným dojmem, a navíc se výrazně lišil od jejich běžného projevu. Při poslechové analýze byl projev 11 mluvčích z 80 vyhodnocen jako relativně přirozeně znějící a jejich identifikovatelnost na základě maskovaného projevu zároveň jako středně obtížná či obtížná. Poslechové testy pak ukázaly, že žádný z maskovaných projevů mluvčích nebyl respondenty identifikován jako spíše či jednoznačně přirozený. S výjimkou maskovaných projevů dvou mluvčích, u kterých respondenti na otázku, zda se jedná o maskování hlasu, odpovídali většinou „těžko říci“, byly maskované projevy všech mluvčích identifikovány jako spíše nebo jednoznačně maskované.

Tyto výsledky napovídají, že maskování hlasu je v drtivé většině případů poměrně snadno odhalitelné.

Druhý percepční test ověřoval obtížnost identifikace mluvčího na základě jeho maskovaného projevu pomocí modelové rekognice, při níž byl posluchačům vždy předložen maskovaný projev mluvčího, který měli spojit s jeho nemaskovaným hlasem, přičemž prováděli rekognici mezi třemi nahrávkami přirozeného projevu: onoho maskujícího mluvčího a dvou figurantů.

Při poslechové analýze bylo 49 mluvčích (tj. 61,25 %) zaznamenáno jakožto snadno identifikovatelných na základě svého maskovaného projevu. Pro detailnější analýzu bylo vybráno 15 mluvčích, jejichž identifikovatelnost byla při poslechové analýze vyhodnocena jako středně obtížná nebo obtížná. Výsledky percepčních testů ukázaly, že nižší než 50% úspěšnost rekognice nastala u 9 mluvčích z 15. Ze vzorku podrobeného percepčním testům tedy relativně efektivní způsob maskování zvolilo 60 % mluvčích, z celkového počtu jednotlivců, kteří se účastnili nahrávání, tvoří těchto 9 mluvčích pouze 11,25 %. Pokud by tento poměr úspěšnosti identifikace byl zachován u všech 31 mluvčích, kteří byli při poslechové analýze vyhodnoceni jako středně obtížně nebo obtížně identifikovatelní, z 80 nahraných mluvčích by svůj hlas natolik efektivně, aby byla úspěšnost rekognice méně než 50 %, maskovalo 18 jedinců, tj. 22,5 % mluvčích.

Zdá se tedy, že schopnost efektivně maskovat svůj hlas má relativně nízký počet mluvčích.

Analýza byla prováděna na základě čteného projevu. Vyšší výpovědní hodnotu o strategiích volených mluvčími při maskování hlasu porovnatelných s forenzní praxí by pravděpodobně měl projev spontánní, ten by však byl náročný na zpracování (kvůli čtenějšímu výskytu dysfluencí, hláskových redukcí apod.) a byl by zapotřebí obsáhlejší materiál vzhledem ke skutečnosti, že by při výzkumu nebylo možné mít kontrolu nad obsahem textu, tj. např. nad výskytem segmentů potřebných k akustické analýze.

Nevýhoda textů předčítaných mluvčími běžným i maskovaným hlasem (viz příloha 1) spočívala mimo jiné v jejich obsahu. V prvním případě, kdy se jednalo o předčítání dialogu „Milánka a maminky“, mohli mít mluvčí tendenci nevědomě modifikovat svůj řečový projev, de facto „hrát role“ postav vystupujících v textu. Mohli také produkovat delší pauzy než při čtení textu s neutrálním obsahem, aby podpořili vyznění předčítané pasáže.

Maskovaným hlasem pak mluvčí předčítali text, který tvořil jednu stranu telefonního rozhovoru dvou „mafíánů“. Tento text mohl být určitým způsobem sugestivní, mluvčí mohl spíše „hrát mafiána“ a tím modifikovat svůj řečový projev jiným způsobem, než kdyby byl instruován, aby libovolným způsobem maskoval svůj hlas při předčítání

neutrálního textu. Skutečnost, že se jednalo o jednu stranu dialogu, mohla také zapříčinit delší pauzy, než jaké by mluvčí používal při četbě neutrálního textu, a tím ovlivnit jeho mluvní tempo (podrobněji viz výše).

Do budoucna by mohlo být vhodné sestavit pro tyto účely text obsahující neutrální, co nejméně expresivní repliky. Pro lepší porovnatelnost vzorků by mohlo být přínosné, kdyby mluvčí četli maskovaným i nemaskovaným hlasem týž text.

Výsledky lépe vypovídající o strategiích, které mluvčí volí při maskování hlasu, jakož i o jejich konzistentnosti v maskování v průběhu času, by mohl přinést výzkum delších spontánních maskovaných projevů, případně nahrávaných opakovaně v delším časovém rozmezí. Takový výzkum by však byl náročný jak časově, tak i co se týče získávání respondentů.

## Reference

- Anolli, L. & Ciceri, R. (1997). The Voice of Deception: Vocal Strategies of Naive and Able Liars. *Journal of Nonverbal Behavior*, 21, pp. 259-284.
- Braun, A. & Rosin, A. (2015). On the Speaker-Specificity of Hesitation Markers. *Proceedings of 18th ICPhS*.
- Clark, J. & Foulkes, P. (2007). Identification of voices in electronically disguised speech. *The International Journal of Speech, Language and the Law*, 14, pp. 195-221.
- Eriksson, A. (2010). The Disguised Voice: Imitating Accents or Speech Styles and Impersonating Individuals. In: Llamas, C. & Watt, D. (Eds.), *Language and Identities*, pp. 86-96. Edinburgh: Edinburgh University Press.
- Eriksson, E., Sullivan, K., Zetterholm, E., Czigler, P., Green, J., Skagerstrand, Å. & van Doorn, J. (2010). Detection of imitated voices: who are reliable earwitnesses? *The International Journal of Speech, Language and the Law*, 17, pp. 25-44.
- Figueiredo, R. M. & Britto, H. S. (1996). A report on the acoustic effects of one type of disguise. *Forensic Linguistics*, 3, pp. 168-175.
- Harnsberger, J. D., Hollien, H., Martin, C. A. & Hollien, K. A. (2009). Stress and Deception in Speech: Evaluating Layered Voice Analysis. *Journal of Forensic Sciences*, 54, pp. 642-650.
- Hirson, A. & Duckworth, M. (1995). Forensic implications of vocal creak as voice disguise. In: *BEIPHOL 64, Studies in Forensic Phonetics*, pp. 67-76.
- Hollien, H. (2002). *Forensic Voice Identification*. London: Academic Press.
- Kirchhübel, C. & Howard, D. M. (2011). Investigating the acoustic characteristics of deceptive speech. *Proceedings of 17th ICPhS*, pp. 1094-1097.
- Künzel, H. J. (2000). Effects of voice disguise on speaking fundamental frequency. *Forensic Linguistics*, 7, pp. 149-179.
- Masthoff, H. (1996). A report on a voice disguise experiment. *Forensic Linguistics*, 3, pp. 160-167.
- Neuhauser, S. & Simpson, A. P. (2007). Phonetic correlates of accent authenticity in voice disguise. *Proceedings of IAFPA 2007*.
- Neuhauser, S. (2008). Voice disguise using a foreign accent: phonetic and linguistic variation. *The International Journal of Speech, Language and the Law*, 15, pp. 131-159.

- Orchard, T. L. & Yarmey, A. D. (1995). The Effects of Whispers, Voice-Sample Duration, and Voice Distinctiveness on Criminal Speaker Identification. *Applied Cognitive Psychology*, 9, pp. 249-260.
- Pollák, P., Volín, J. a Skarnitzl, R. (2007). HMM-based phonetic segmentation in Praat environment. In: Proceedings of XIIth “Speech and computer – SPECOM 2007”, 537–541.
- Skarnitzl, R. (2014). *Fonetická identifikace mluvčího*. Praha: Univerzita Karlova v Praze, Filozofická fakulta
- Zetterholm, E. (2007). Detection of Speaker Characteristics Using Voice Imitation. In: Miller, C. (Ed.), *Speaker Classification II, LNAI 4441*, pp. 192-205.

## Přílohy

### Příloha 1 – předčítané texty

Nemaskovaným hlasem mluvčí předčítali následující text:

Milánek

Maminka se zeptala Milana: „Milánku, už máš hotový úkol? Kdy ho budeš psát?“

Milan chvilku přemýšlel a pak odpověděl: „Já musím napsat pár souvětí na Říhovou, kde budou nějaké gramatické figle. Například vztažné věty, čárky před „a“ a podobně. *Vlastně jsem tě chtěl poprosit, jestli mi s tím nepomůžeš.*“

„*Můžeme se na to mrknout klidně hned,*“ řekla maminka, „jen bych dala vařit vodu na čaj a podívám se, jestli máme citrony. Jak dlouho nám ten úkol zabere? Bude to těžké?“

„No, mají tam být i různé příklady na zastaralou a knižní slovní zásobu. Skoro půlku jsem už ve škole udělal, ale *moc dobře mi to nešlo. Chtěl bych začít co nejdřív.* Až budu hotov, došel bych ti do lékárny pro ten pneumocyt. A potom půjdu hrát fotbal, včera jsem dal čtyři góly. Nebýt Láďových faulů, mohlo jich být víc. Jen míč budu muset přifouknout. Neboj se, dám pozor na auta.“

Maskovaným hlasem mluvčí četli tento text:

Milánku, seš to ty? Zdar, tady Viktor. Je vzduch čistej? Můžu mluvit?

Jenom volám, že Říhovou jsem včera už ve škole udělal. Trochu se vztekala, *moc dobře mi to nešlo*, ale nakonec kápala božškou. Teď ještě zmáčknout ředitele a budeme mít všechno. Pak to rozjedem.

*Vlastně jsem tě chtěl poprosit, jestli mi s tím nepomůžeš.*

Super, díky. Už máš hotovej ten svůj úkol?

Jasně, to víš, že ti taky píchnu. *Můžeme se na to mrknout klidně hned.* Jak dlouho myslíš, že nám to zabere?

No moc dlouho bych nečekal, *chtěl bych začít co nejdřív.* Podívám se, jestli máme nachystaný bouchačky a pak mrknu na auta. Myslíš, že to bude těžký?

Neboj se, dám pozor, aby se to neprovalilo. Až budu hotovej, tak ti brnknu. A potom půjdu na zaslouženého panáka.

Tak jo, Milánku, domluveno. Měj se.

Fráze, které se objevují v obou textech, jsou vyznačeny kurzívou. Tyto fráze byly využity při percepčních testech.

**Příloha 2 – výsledky akustické analýzy vokálních formantů při maskovaných a nemaskovaných projevech mluvčích a vyčíslení p-hodnoty pomocí dvouvýběrového t-testu**

i				
mluvčí	přirozený	maskovaný	rozdíl	t-test
DRID	425,13	448,62	23,49	0,004
JELS	404,47	468,42	63,95	<0,05
KAND	392,84	415,54	22,69	<0,05
KATR	326,61	299,07	-27,54	<0,05
KREM	419,83	390,91	-28,92	0,03
KUCK	407,27	437,18	29,91	0,03
MAKN	430,50	719,97	289,47	<0,05
MAND	389,62	420,93	31,31	<0,05
OLET	405,36	339,55	-65,81	<0,05
ROLK	394,86	407,66	12,80	0,44
SJAN	398,37	410,18	11,82	0,11
VEDR	429,67	419,81	-9,86	0,16
VLAJ	419,31	353,41	-65,90	<0,05
ZALT	417,68	448,67	30,99	<0,05
ZENT	423,33	371,10	-52,23	<0,05

**Tabulka 7-F1-1:** rozdíl F1 vokálu /i/ při maskovaném a nemaskovaném projevu

e				
mluvčí	přirozený	maskovaný	rozdíl	t-test
DRID	499,08	573,00	73,92	<0,05
JELS	488,50	621,44	132,94	<0,05
KAND	422,02	502,14	80,12	<0,05
KATR	451,00	338,62	-112,38	<0,05
KREM	471,03	555,10	84,07	<0,05
KUCK	456,65	625,27	168,63	<0,05
MAKN	489,67	801,20	311,53	<0,05
MAND	467,23	583,93	116,70	<0,05
OLET	436,44	382,22	-54,22	<0,05
ROLK	517,98	567,49	49,51	<0,05
SJAN	441,72	500,32	58,60	<0,05
VEDR	492,09	554,44	62,35	<0,05
VLAJ	522,11	371,20	-150,90	<0,05
ZALT	471,00	531,48	60,48	<0,05
ZENT	486,63	445,74	-40,89	<0,05

**Tabulka 7-F1-2:** rozdíl F1 vokálu /e/ při maskovaném a nemaskovaném projevu

a				
mluvčí	přirozený	maskovaný	rozdíl	t-test
DRID	638,84	643,24	4,40	0,54
JELS	593,56	755,87	162,31	<0,05
KAND	525,64	682,87	157,24	<0,05
KATR	637,35	409,29	-228,05	<0,05
KREM	602,70	724,75	122,05	<0,05
KUCK	577,19	884,15	306,96	<0,05
MAKN	600,88	954,37	353,48	<0,05
MAND	656,34	837,63	181,29	<0,05
OLET	540,97	480,38	-60,60	<0,05
ROLK	669,50	667,77	-1,73	0,92
SJAN	550,50	594,16	43,66	<0,05
VEDR	589,54	627,24	37,71	<0,05
VLAJ	522,11	371,20	-150,90	<0,05
ZALT	624,91	679,49	54,58	<0,05
ZENT	582,85	587,10	4,26	0,70

**Tabulka 7-F1-3:** rozdíl F1 vokálu /a/ při maskovaném a nemaskovaném projevu

o				
mluvčí	přirozený	maskovaný	rozdíl	t-test
DRID	479,70	526,14	46,44	0,92
JELS	473,62	636,74	163,12	<0,05
KAND	434,83	579,35	144,52	<0,05
KATR	394,03	342,22	-51,82	<0,05
KREM	459,36	508,13	48,77	<0,05
KUCK	458,60	696,49	237,89	<0,05
MAKN	675,27	852,21	176,95	<0,05
MAND	463,48	588,94	125,47	<0,05
OLET	443,85	398,28	-45,57	<0,05
ROLK	500,97	513,07	12,10	0,40
SJAN	426,59	526,54	99,95	<0,05
VEDR	455,02	550,21	95,19	<0,05
VLAJ	496,53	399,36	-97,17	<0,05
ZALT	456,19	507,06	50,87	<0,05
ZENT	478,07	452,20	-25,87	<0,05

**Tabulka 7-F1-4:** rozdíl F1 vokálu /o/ při maskovaném a nemaskovaném projevu



u				
mluvci	přirozený	maskovaný	rozdíl	t-test
DRID	352,87	370,53	17,66	0,03
JELS	382,93	539,48	156,55	<0,05
KAND	373,61	443,63	70,02	<0,05
KATR	296,32	291,52	-4,80	0,40
KREM	438,26	419,15	-19,11	0,48
KUCK	399,75	586,97	187,22	<0,05
MAKN	428,51	885,09	456,58	<0,05
MAND	349,19	436,12	86,92	<0,05
OLET	396,22	329,17	-67,06	0,02
ROLK	391,17	482,85	91,68	<0,05
SJAN	370,16	466,25	96,09	<0,05
VEDR	381,63	452,62	70,99	<0,05
VLAJ	381,66	329,44	-52,22	<0,05
ZALT	378,69	417,90	39,20	<0,05
ZENT	381,62	404,52	22,91	0,14

**Tabulka 7-F1-5:** rozdíl F1 vokálu /u/ při maskovaném a nemaskovaném projevu

i:				
mluvčí	přirozený	maskovaný	rozdíl	t-test
DRID	338,84	317,45	-21,39	0,015
JELS	370,84	366,68	-4,16	0,64
KAND	304,20	348,74	44,54	<0,05
KATR	290,11	259,82	-30,30	<0,05
KREM	311,14	275,44	-35,70	<0,05
KUCK	343,43	538,34	194,91	<0,05
MAKN	337,14	687,28	350,14	<0,05
MAND	308,84	307,24	-1,60	0,91
OLET	321,82	297,45	-24,37	<0,05
ROLK	317,22	336,17	18,94	0,41
SJAN	326,26	354,76	28,50	<0,05
VEDR	336,37	367,19	30,82	<0,05
VLAJ	336,59	289,35	-47,24	<0,05
ZALT	296,60	338,01	41,41	<0,05
ZENT	367,04	348,46	-18,58	0,10

**Tabulka 7-F1-6:** rozdíl F1 vokálu /i:/ při maskovaném a nemaskovaném projevu

a:				
mluvčí	přirozený	maskovaný	rozdíl	t-test
DRID	638,84	643,24	4,40	0,54
JELS	637,47	805,67	168,20	<0,05
KAND	591,12	738,42	147,29	<0,05
KATR	695,57	478,68	-216,89	<0,05
KREM	698,95	784,99	86,04	<0,05
KUCK	619,13	938,05	318,92	<0,05
MAKN	686,55	1058,28	371,72	<0,05
MAND	773,63	914,06	140,43	<0,05
OLET	614,82	509,21	-105,61	0,02
ROLK	718,98	805,02	86,04	<0,05
SJAN	621,45	644,20	22,76	0,06
VEDR	634,64	659,33	24,69	<0,05
VLAJ	711,15	653,79	-57,36	0,05
ZALT	657,08	684,32	27,24	<0,05
ZENT	627,72	608,63	-19,09	0,09

**Tabulka 7-F1-7:** rozdíl F1 vokálu /a:/ při maskovaném a nemaskovaném projevu

u:				
mluvčí	přirozený	maskovaný	rozdíl	t-test
DRID	324,47	369,90	45,43	0,04
JELS	358,24	488,95	130,71	<0,05
KAND	343,07	419,38	76,31	<0,05
KATR	306,70	280,93	-25,78	<0,05
KREM	386,42	363,38	-23,04	0,58
KUCK	368,75	553,23	184,48	<0,05
MAKN	366,12	770,54	404,42	<0,05
MAND	319,08	356,48	37,41	<0,05
OLET	349,08	308,35	-40,73	<0,05
ROLK	366,19	335,59	-30,60	<0,05
SJAN	358,48	435,58	77,09	<0,05
VEDR	353,46	387,95	34,49	<0,05
VLAJ	387,64	343,28	-44,37	0,21
ZALT	323,99	393,47	69,48	<0,05
ZENT	336,47	375,70	39,23	0,03

**Tabulka 7-F1-8:** rozdíl F1 vokálu /u:/ při maskovaném a nemaskovaném projevu

mluvci	i			
	přirozený	maskovaný	rozdíl	t-test
DRID	1645,55	1389,16	-256,39	<0,05
JELS	1272,72	1279,41	6,69	0,88
KAND	1651,22	1898,55	247,32	<0,05
KATR	1660,82	1475,02	-185,80	<0,05
KREM	1406,67	1460,86	54,19	0,21
KUCK	1669,19	1603,26	-65,93	0,12
MAKN	1183,69	1470,48	286,79	<0,05
MAND	1619,07	1667,12	48,05	0,23
OLET	1741,79	1376,72	-365,07	<0,05
ROLK	1302,81	1348,07	45,26	0,41
SJAN	1548,22	1665,07	116,85	<0,05
VEDR	1632,70	1444,64	-188,05	<0,05
VLAJ	1466,98	1670,81	203,83	<0,05
ZALT	1521,94	1359,22	-162,73	<0,05
ZENT	423,33	371,10	-52,23	<0,05

Tabulka 7-F2-1: rozdíl F2 vokálu /i/ při maskovaném a nemaskovaném projevu

mluvci	e			
	přirozený	maskovaný	rozdíl	t-test
DRID	1614,45	1640,01	25,56	0,19
JELS	1662,29	1329,40	-332,88	<0,05
KAND	1756,22	1784,97	28,75	0,38
KATR	1540,79	1692,29	151,50	<0,05
KREM	1544,33	1750,08	205,75	<0,05
KUCK	1633,33	1744,43	111,11	<0,05
MAKN	1481,61	1646,30	164,69	<0,05
MAND	1672,31	1880,17	207,86	<0,05
OLET	1667,68	1302,60	-365,08	<0,05
ROLK	1689,74	1875,79	186,05	<0,05
SJAN	1496,92	1643,70	146,78	<0,05
VEDR	1590,91	1607,79	16,88	0,41
VLAJ	1716,53	1552,69	-163,84	<0,05
ZALT	1656,64	1742,35	85,71	<0,05
ZENT	1428,50	1599,58	171,08	<0,05

Tabulka 7-F2-2: rozdíl F2 vokálu /e/ při maskovaném a nemaskovaném projevu

mluvci	a			
	přirozený	maskovaný	rozdíl	t-test
DRID	1526,79	1397,92	-128,87	<0,05
JELS	1445,06	1329,65	-115,40	<0,05
KAND	1760,23	1760,32	0,10	1,00
KATR	1483,46	1779,45	295,99	<0,05
KREM	1386,16	1544,99	158,83	<0,05
KUCK	1438,64	1601,42	162,78	<0,05
MAKN	1250,65	1531,30	280,65	<0,05
MAND	1552,36	1639,31	86,95	<0,05
OLET	1560,38	1322,32	-238,06	<0,05
ROLK	1417,52	1540,67	123,15	<0,05
SJAN	1462,27	1526,51	64,24	<0,05
VEDR	1354,26	1322,70	-31,56	0,35
VLAJ	1444,18	1554,25	110,07	<0,05
ZALT	1417,57	1498,78	81,21	<0,05
ZENT	1303,79	1591,05	287,25	<0,05

Tabulka 7-F2-3: rozdíl F2 vokálu /a/ při maskovaném a nemaskovaném projevu

mluvci	o			
	přirozený	maskovaný	rozdíl	t-test
DRID	1035,65	1034,30	-1,35	0,92
JELS	1054,95	1094,33	39,38	<0,05
KAND	1078,13	1052,47	-25,66	0,46
KATR	940,50	1236,56	296,05	<0,05
KREM	1110,83	1142,79	31,95	0,10
KUCK	979,69	1339,57	359,87	<0,05
MAKN	1185,83	1294,66	108,84	<0,05
MAND	1138,68	1146,92	8,24	0,63
OLET	1068,56	1061,47	-7,09	0,70
ROLK	1179,95	1222,04	42,10	0,05
SJAN	1087,87	1321,05	233,18	<0,05
VEDR	933,33	1055,69	122,36	<0,05
VLAJ	1170,62	1221,72	51,10	0,03
ZALT	1086,79	1138,07	51,27	<0,05
ZENT	1002,84	1138,01	135,17	<0,05

Tabulka 7-F2-4: rozdíl F2 vokálu /o/ při maskovaném a nemaskovaném projevu

u				
mluvčí	přirozený	maskovaný	rozdíl	t-test
DRID	905,29	910,61	5,32	0,86
JELS	973,11	1082,78	109,67	<0,05
KAND	973,82	815,92	-157,91	<0,05
KATR	845,57	789,97	-55,59	0,03
KREM	1147,69	1124,97	-22,72	0,64
KUCK	970,72	1331,66	360,93	<0,05
MAKN	1039,21	1239,79	200,59	<0,05
MAND	1051,28	981,28	-70,00	0,02
OLET	1101,17	963,92	-137,25	<0,05
ROLK	1087,83	1199,64	111,81	<0,05
SJAN	1060,31	1287,69	227,37	<0,05
VEDR	840,08	1009,80	169,72	<0,05
VLAJ	1006,20	1220,81	214,61	<0,05
ZALT	1077,82	959,86	-117,96	<0,05
ZENT	879,86	1131,39	251,53	<0,05

**Tabulka 7-F2-5:** rozdíl F1 vokálu /u/ při maskovaném a nemaskovaném projevu

i:				
mluvčí	přirozený	maskovaný	rozdíl	t-test
DRID	1962,59	1364,25	-598,34	<0,05
JELS	1656,97	1306,21	-350,76	<0,05
KAND	1673,41	1705,56	32,15	0,65
KATR	1778,68	1677,78	-100,90	0,52
KREM	1678,64	1408,61	-270,02	<0,05
KUCK	1892,62	1754,22	-138,40	0,01
MAKN	1316,41	1657,54	341,13	<0,05
MAND	1665,67	1776,17	110,49	0,04
OLET	1607,82	1458,65	-149,17	<0,05
ROLK	1247,92	1443,65	195,72	<0,05
SJAN	1611,30	1695,83	84,53	0,19
VEDR	1585,68	1562,98	-22,69	0,79
VLAJ	1623,43	1844,61	221,18	<0,05
ZALT	1668,05	1582,52	-85,53	0,36
ZENT	1854,70	1874,62	19,92	0,71

**Tabulka 7-F2-6:** rozdíl F2 vokálu /i:/ při maskovaném a nemaskovaném projevu

a:				
mluvčí	přirozený	maskovaný	rozdíl	t-test
DRID	1451,74	1371,52	-80,22	0,01
JELS	1345,59	1405,00	59,41	<0,05
KAND	1695,11	1694,29	-0,82	0,98
KATR	1484,36	1727,37	243,01	<0,05
KREM	1372,23	1601,63	229,40	<0,05
KUCK	1334,87	1513,45	178,59	<0,05
MAKN	1194,63	1496,83	302,20	<0,05
MAND	1493,48	1533,48	40,00	<0,05
OLET	1421,15	1368,43	-52,72	0,14
ROLK	1364,53	1439,61	75,08	<0,05
SJAN	1430,22	1526,20	95,98	<0,05
VEDR	1174,68	1464,03	289,34	<0,05
VLAJ	1386,99	1454,20	67,22	0,22
ZALT	1425,10	1470,52	45,42	0,03
ZENT	1314,11	1375,55	61,45	0,05

**Tabulka 7-F2-7:** rozdíl F2 vokálu /a:/ při maskovaném a nemaskovaném projevu

u:				
mluvčí	přirozený	maskovaný	rozdíl	t-test
DRID	857,77	952,16	94,39	0,15
JELS	877,20	1120,36	243,16	<0,05
KAND	1233,81	739,18	-494,63	<0,05
KATR	872,09	732,69	-139,40	<0,05
KREM	1090,04	1125,40	35,37	0,56
KUCK	898,10	1213,99	315,89	<0,05
MAKN	904,76	1077,60	172,84	<0,05
MAND	976,64	840,45	-136,20	<0,05
OLET	889,96	999,47	109,51	<0,05
ROLK	1176,68	961,23	-215,45	<0,05
SJAN	971,74	1169,85	198,10	<0,05
VEDR	831,45	1033,51	202,06	0,02
VLAJ	1116,75	1273,17	156,42	<0,05
ZALT	945,33	1001,52	56,19	0,13
ZENT	885,41	1134,58	249,18	<0,05

**Tabulka 7-F2-8:** rozdíl F2 vokálu /u:/ při maskovaném a nemaskovaném projevu



i				
mluvci	přirozený	maskovaný	rozdíl	t-test
DRID	2408,35	2137,01	-271,34	<0,05
JELS	2148,43	1990,29	-158,14	<0,05
KAND	2376,71	2555,66	178,95	<0,05
KATR	2374,94	2210,99	-163,95	<0,05
KREM	2044,66	2234,43	189,77	<0,05
KUCK	2383,58	2257,52	-126,06	<0,05
MAKN	1947,21	2054,54	107,33	<0,05
MAND	2241,25	2342,79	101,54	<0,05
OLET	2384,18	2032,67	-351,51	<0,05
ROLK	2244,59	2216,44	-28,16	0,45
SJAN	2329,74	2364,22	34,47	0,29
VEDR	2384,52	2351,77	-32,75	0,24
VLAJ	2303,96	2375,80	71,85	0,03
ZALT	2376,60	2285,55	-91,05	0,04
ZENT	2346,36	2419,96	73,60	<0,05

**Tabulka 7-F3-1:** rozdíl F3 vokálu /i/ při maskovaném a nemaskovaném projevu

e				
mluvci	přirozený	maskovaný	rozdíl	t-test
DRID	2562,81	2500,79	-62,02	<0,05
JELS	2522,72	2443,12	-79,60	<0,05
KAND	2724,40	2329,43	-394,97	<0,05
KATR	2455,76	2477,42	21,66	0,24
KREM	2549,62	2760,66	211,04	<0,05
KUCK	2488,37	2655,93	167,56	<0,05
MAKN	2381,75	2518,08	136,33	<0,05
MAND	2546,45	2538,61	-7,84	0,67
OLET	2609,90	2188,14	-421,76	<0,05
ROLK	2534,34	2660,61	126,27	<0,05
SJAN	2503,32	2604,96	101,64	<0,05
VEDR	2488,20	2886,77	398,57	<0,05
VLAJ	2583,72	2481,74	-101,98	<0,05
ZALT	2663,91	2627,47	-36,44	0,10
ZENT	2401,62	2574,92	173,31	<0,05

**Tabulka 7-F3-2:** rozdíl F3 vokálu /e/ při maskovaném a nemaskovaném projevu

a				
mluvci	přirozený	maskovaný	rozdíl	t-test
DRID	3254,23	2820,65	-433,58	<0,05
JELS	2908,88	2390,73	-518,15	<0,05
KAND	3333,47	3542,51	209,04	<0,05
KATR	3118,99	2794,26	-324,73	<0,05
KREM	2996,35	3045,09	48,74	0,11
KUCK	2726,86	3120,34	393,47	<0,05
MAKN	2717,80	2817,58	99,79	<0,05
MAND	2982,77	2720,41	-262,36	<0,05
OLET	3197,21	2465,39	-731,82	<0,05
ROLK	2629,85	2832,91	203,06	<0,05
SJAN	2836,24	3065,90	229,65	<0,05
VEDR	2810,85	3191,22	380,37	<0,05
VLAJ	2755,76	3014,56	258,80	<0,05
ZALT	2952,68	2732,32	-220,36	<0,05
ZENT	2806,18	3032,41	226,23	<0,05

**Tabulka 7-F3-3:** rozdíl F1 vokálu /a/ při maskovaném a nemaskovaném projevu

o				
mluvci	přirozený	maskovaný	rozdíl	t-test
DRID	2479,63	2335,13	-144,50	<0,05
JELS	2173,87	2117,25	-56,62	<0,05
KAND	2454,94	2009,72	-445,22	<0,05
KATR	2376,33	2367,31	-9,02	0,66
KREM	2584,68	2714,09	129,41	<0,05
KUCK	2436,43	2611,99	175,55	<0,05
MAKN	2405,84	2465,28	59,44	<0,05
MAND	2437,97	2442,19	4,21	0,82
OLET	2358,06	2159,48	-198,59	<0,05
ROLK	2464,70	2562,96	98,26	<0,05
SJAN	2369,92	2251,25	-118,67	<0,05
VEDR	2397,27	2873,71	476,43	<0,05
VLAJ	2611,16	2486,77	-124,39	<0,05
ZALT	2467,39	2434,41	-32,98	0,12
ZENT	2446,77	2572,15	125,38	<0,05

**Tabulka 7-F3-4:** rozdíl F3 vokálu /o/ při maskovaném a nemaskovaném projevu

u				
mluvčí	přirozený	maskovan	rozdíl	t-test
DRID	2479,63	2335,13	-144,50	<0,05
JELS	2083,41	1735,70	-347,71	<0,05
KAND	2261,14	2024,54	-236,60	<0,05
KATR	2160,38	1965,90	-194,48	<0,05
KREM	2330,54	1918,25	-412,29	<0,05
KUCK	2245,08	2195,33	-49,75	0,30
MAKN	1956,05	2225,32	269,27	<0,05
MAND	2130,93	2098,92	-32,02	0,44
OLET	2124,99	1829,44	-295,55	<0,05
ROLK	2239,53	2148,73	-90,81	0,13
SJAN	2188,80	2104,91	-83,89	0,03
VEDR	2172,71	2209,96	37,25	0,37
VLAJ	2091,88	2194,31	102,43	0,06
ZALT	2220,79	1962,73	-258,06	<0,05
ZENT	2211,95	2239,71	27,76	0,52

**Tabulka 7-F3-5:** rozdíl F3 vokálu /u/ při maskovaném a nemaskovaném projevu

i:				
mluvci	přirozený	maskovaný	rozdíl	t-test
DRID	2671,48	2411,30	-260,17	<0,05
JELS	2506,48	2100,79	-405,69	<0,05
KAND	2523,46	2609,38	85,92	<0,05
KATR	2604,56	2444,10	-160,46	<0,05
KREM	2377,24	2449,82	72,58	0,05
KUCK	2529,01	2563,35	34,33	0,57
MAKN	2262,85	2420,59	157,73	<0,05
MAND	2424,15	2440,11	15,95	0,91
OLET	2663,46	2227,18	-436,28	<0,05
ROLK	2435,97	2437,59	1,62	0,96
SJAN	2453,13	2455,40	2,27	0,96
VEDR	2460,74	2557,39	96,65	<0,05
VLAJ	2596,95	2548,02	-48,93	0,17
ZALT	2557,79	2547,03	-10,76	0,81
ZENT	2578,34	2483,45	-94,88	<0,05

**Tabulka 7-F3-6:** rozdíl F3 vokálu /i:/ při maskovaném a nemaskovaném projevu

a:				
mluvčí	přirozený	maskovaný	rozdíl	t-test
DRID	3168,93	2914,60	-254,33	<0,05
JELS	2908,27	2408,68	-499,60	<0,05
KAND	3306,66	3341,75	35,09	0,64
KATR	3219,40	2856,50	-362,91	<0,05
KREM	2908,01	3118,77	210,76	<0,05
KUCK	2673,53	3160,92	487,39	<0,05
MAKN	2776,81	2831,76	54,94	0,16
MAND	2847,96	2831,12	-16,84	0,57
OLET	2850,33	2536,22	-314,11	<0,05
ROLK	2615,45	2843,51	228,06	<0,05
SJAN	2771,94	3129,13	357,19	<0,05
VEDR	2568,39	3143,55	575,16	<0,05
VLAJ	2561,83	2944,10	382,27	<0,05
ZALT	2805,87	2775,16	-30,71	0,47
ZENT	2814,91	2769,14	-45,78	0,31

**Tabulka 7-F3-7:** rozdíl F3 vokálu /a:/ při maskovaném a nemaskovaném projevu

u:				
mluvčí	přirozený	maskovaný	rozdíl	t-test
DRID	2233,07	1904,28	-328,79	<0,05
JELS	2078,55	1676,09	-402,46	<0,05
KAND	2352,92	1981,63	-371,30	<0,05
KATR	2158,13	1875,47	-282,65	<0,05
KREM	2188,11	2240,44	52,33	0,48
KUCK	2110,14	2077,23	-32,91	0,61
MAKN	2111,33	2256,34	145,01	<0,05
MAND	2181,44	1979,96	-201,48	<0,05
OLET	2159,40	1719,82	-439,59	<0,05
ROLK	2192,98	2084,74	-108,24	0,30
SJAN	2226,15	1905,05	-321,10	<0,05
VEDR	2282,61	2182,48	-100,13	0,22
VLAJ	2172,42	2170,27	-2,15	0,98
ZALT	2171,60	2151,41	-20,18	0,77
ZENT	2290,38	2342,15	51,78	0,32

**Tabulka 7-F3-8:** rozdíl F3 vokálu /u:/ při maskovaném a nemaskovaném projevu